Ulam Quarterly — Volume 3, Number 2, 1996

Number of Left-Product Coefficients in a Mendelian Stochastic Algebra

W.A. Beyer

Theoretical Division
Los Alamos National Laboratory
Los Alamos, NM 87545

and

L.M. Holm

Los Alamos, NM 87545

Abstract

We consider the number of ratios of three zygotic types in a population of single characteristic diploid mating individuals of three zygotic types: dominant, mixed, and recessive (d, m, and r). The rules of reproduction are the usual ones in Mendelian inheritance: $d + d \rightarrow d$, $d + m \rightarrow d$ and m in equal ratios, $d + r \rightarrow m$, $m + m \rightarrow$ all three in the 1:2:1 ratios, $m + r \rightarrow$ m and r in equal ratios, and finally $r{+}r \rightarrow r$. One starts with a population of size three at generation 0, one each of the three zygotic types. Each member then mates once with each of the three zygotic types from outside the tree of generations and produces three offspring in generation 1. This process is continued, generation after generation. We obtain by use of a multiplication table, ordered triplets that can be interpreted, when normalized, as the probability that a given left- product of d's, m's, and r's is of type d, m, or r. The ordered triplets present in generation n are partitions of 2^{2n}. Distinct triplets give distinct ratios. We calculate the number of triplets in each generation. We show that the number of distinct triplets at the n-th generation is $3 \cdot 2^n$. This is much smaller, for large n, than $3 \cdot 3^n$, the number of elements in each generation. The number of unordered triplets is $2^n + 1$.

1. Introduction.

In this paper we consider the number of ratios of three zygotic types in a population of single characteristic diploid mating individuals of three zygotic types: dominant, mixed, and recessive (d, m, and r). The rules of reproduction are the usual ones in Mendelian inheritance: $d + d \rightarrow d$, $d + m \rightarrow d$ and m in equal ratios, $d + r \rightarrow m$, $m + m \rightarrow$ all three in the 1:2:1 ratios, $m + r \rightarrow m$ and r in equal ratios, and finally $r + r \rightarrow r$.

Suppose one starts with a population of size three at generation 0, one each of the three zygotic types. (One could define a special zygotic type to be used only once. This element will become the multiplicative identity: e. In that case, one could extend the tree backward to generation -1 and make the tree a rooted tree by hypothesizing the element e at generation -1.) Each member then mates once with each of the three zygotic types from outside the tree of generations and produces three offspring in generation 1. Thus there are nine members in generation 1: d^2, dm, dr, md, m^2, mr, rd, rm and r^2. This process is continued, generation after generation. As explained in §2, we obtain by use of the multiplication table (1) below, ordered triplets that can be interpreted, when normalized, as the probability that a given left-product of d's, m's, and r's is of type d, m, or r. The ordered triplets that are present in generation n are partitions of 2^{2n}. Therefore distinct triplets give distinct ratios. Hence we use the terms triplets and ratios interchangeably. We are interested in the number of possible ratios or triplets present in each generation. We show in this paper that the number of distinct ratios at the n-th generation is $3 \cdot 2^n$. This is much smaller, for large n, than $3 \cdot 3^n$, the number of elements in each generation. If one counts unordered ratios, then the number of distinct ratios is $2^n + 1$.

In discussing the meaning of these results, one needs to distinguish the codominant (a distinct phenotype for each genotype) situation from the noncodominant situation. Suppose we consider a codominant gene and that we know the zygotic type of a given individual and the zygotic types of its ancestors for n generations back. Then given the zygotic type of the individual, one can calculate the probabilities of the three types without reference to the ancestors. The ratios, for a given sequence of n ancestors, tell us what the probability distribution is for the three types in the nth generation. Thus the sets of triplets obtained tell us the distributions of probabilities. If the gene is noncodominant, one can cannot determine the sequence of ancestors and therefore which ratio applies. What the counting of ratios does for us is to put limits on the amount of evolution possible in this very restricted model.

One of the shortcomings of this model is that mating by objects of the tree is always done with objects of known type coming from outside the tree. This is similar to a patriarchal family tree in which only males are listed. Nonfamilial females are imported and familial females are exported. This might partly account for the few number of ratios we find. It is hoped in the future to carry out a project in which we make a more complete counting of possible ancestral trees. Such a project would involve the counting of non-associative combinations. See [3].

The proof in outline goes as follows. Representation of the evolution of the population by a rooted trees shows there cannot be more than $3 \cdot 3^n$ ratios at the n-th generation. For ordered triplets, the number of ratios is shown to be $3 \cdot 2^n$. For unordered triplets, the number is shown to be $2^n + 1$. The problem is formulated and solved by methods of genetic algebra.

2. Zygotic, Stochastic, and Baric Algebras.

The collection of nonassociative algebras arising in theoretic genetics are called genetic algebras. For a through discussion of them, see Wörz-Busekros [6]. An introduction is given in Etherington [2]. Computational aspects of them are discussed in the manual for the symbolic manipulation system AX-IOM [5]. We consider a commutative, but nonassociative, zygotic algebra $\mathbf{M}$ over the rationals with three basis elements: $d,\ m,\ r$. We define $\mathbf{M}$ by the following multiplication table, giving second degree products as linear combinations of the quantities $d,\ m,\ r$. Multiplication is assumed to be commutative.

$$\begin{aligned} d^2 &= 4d, \\ dm &= 2d + 2m, \\ dr &= 4m, \\ m^2 &= d + 2m + r, \\ mr &= 2m + 2r, \\ r^2 &= 4r. \end{aligned} \tag{1}$$

Multiplication is assumed to take place from the left. We are interested in the number of different ordered triplets $(\delta,\ \mu,\ \rho)$ that arise in expansions of arbitrary left-products of $d,\ m,\ r$, into quantities of the form

$$\delta d + \mu m + \rho r, \tag{2}$$

using the multiplication table (1). The coefficients in (2) are of course nonnegative integers.

When the substitution

$$d = 4\bar{d},\ m = 4\bar{m},\ r = 4\bar{r} \tag{2a}$$

is made, (1) becomes

$$\begin{aligned}
\bar{d}^2 &= \bar{d}, \\
\bar{d}\bar{m} &= \frac{1}{2}\bar{d} + \frac{1}{2}\bar{m}, \\
\bar{d}\bar{r} &= \bar{m}, \\
\bar{m}^2 &= \frac{1}{4}\bar{d} + \frac{1}{2}\bar{m} + \frac{1}{4}\bar{r}, \\
\bar{m}\bar{r} &= \frac{1}{2}\bar{m} + \frac{1}{2}\bar{r}, \\
\bar{r}^2 &= \bar{r},
\end{aligned} \tag{2b}$$

where again multiplication is commutative. We call this algebra $\bar{\mathbf{M}}$. If one sets $a_1 = \bar{d}$, $a_2 = \bar{m}$, and $a_3 = \bar{r}$, then (2b) can be written in the from

$$a_i a_j = \sum_{k=1}^{3} \gamma_{ijk} a_k, \; i, j = 1, 2, 3, \tag{2c}$$

with

$$\sum_{k=1}^{3} \gamma_{ijk} = 1, \quad \gamma_{ijk} \geq 0, \tag{2d}$$

where γ_{ijk} are obtained from (2b). Condition (2d) insures that all products of degree n can be represented in the form of a linear combination of the basis vectors a_1, a_2, a_3 with coefficients that are nonnegative real numbers with sum equal to 1. We write this left-product in the form

$$\text{product} = \sum_{k=1}^{3} \gamma^{[n]}_{\text{product},k} a_k,$$

where

$$\sum_{k=1}^{3} \gamma^{[n]}_{\text{product},k} = 1.$$

The quantity $\gamma^{[n]}_{\text{product},k}$ is the probability that mating of zygotes with a zygote of known type that make up the n-fold left-product produce a zygote of type k, where k is one of d, m, or r. Rather than calculating $\gamma^{[n]}_{\text{product},k}$ for all products of degree n, we find it more convenient to work with triplets of integers.

$\bar{\mathbf{M}}$ is a stochastic algebra with a genetic realization. The basis $(\bar{d}, \bar{m}, \bar{r})$ is called a natural basis. For definitions, see page 12 of [6]. It turns out that any stochastic algebra is baric and can have a norm defined so that it is also a Banach algebra. Note that $\bar{\mathbf{M}}$ does not have a multiplicative identity. As is done in Hille and Phillips, page 118 of [4], it is possible to embed $\bar{\mathbf{M}}$

in a larger algebra $\bar{\bar{\mathbf{M}}}$ that does have a multiplicative identity so that $\bar{\mathbf{M}}$ is isomorphic to a subalgebra of $\bar{\bar{\mathbf{M}}}$.

3. Counting Triplets for Small n.

We have carried out the calculations of the types of triplets appearing in the first four generations using the symbol manipulation program MACSYMA. At the nth generation the types of unordered triplets present are determined by first forming all the left-products that represent the ancestry of each possible zygote present by the following left-product:

$$(d^{t^1_d}(m^{t^1_m}(r^{t^1_r}(d^{t^2_d}(m^{t^2_m}(r^{t^2_r}(\dots(d^{t^l_d}(m^{t^l_m}(r^{t^l_r})))\dots))))))), \tag{3}$$

where the t's are nonnegative integers whose sum is n. Thus the nonzero t's in (3) are compositions of n. See Andrews [1]. We reduce each quantity from (3) into the form (2) using the multiplication table. To obtain the unordered triplets present in the $(n+1)$th generation, we successively multiple the resulting linear forms on the left by d, m, and r to obtain second degree left-products. The multiplication table is then used to reduce these sums of second degree products to linear forms.

We then count the number of distinct unordered triplets, or equivalently the distinct zygotic ratios. We arrange entries in the ordered triplets in descending order to give unordered triplets and obtain the results given in the tables below. The form $\{[a,b,c];e\}$ denotes the unordered triplet $[a,b,c]$ with multiplicity e, where multiplicity denotes the number of ordered triplets corresponding to the given unordered triplet. Square brackets denote unordered triplets, round brackets denote ordered triplets.

Table of Ordered Triplets

Generation #	Ordered triplets	# of ordered triplets
0	$(0,0,1)$ $(0,1,0)$ $(1,0,0)$	$3 = 3\cdot 2^0$
1	$(0,0,4)$ $(0,4,0)$ $(4,0,0)$ $(0,2,2)$ $(1,2,1)$ $(2,2,0)$	$6 = 3\cdot 2^1$
2	$(0,0,16)$ $(0,16,0)$ $(16,0,0)$ $(0,4,12)$ $(4,8,4)$ $(12,4,0)$ $(0,8,8)$ $(6,8,2)$ $(8,8,0)$ $(0,12,4)$ $(2,8,6)$ $(4,12,0)$	$12 = 3\cdot 2^2$
3	$(0,0,64)$ $(0,64,0)$ $(64,0,0)$ $(0,8,56)$ $(28,32,4)$ $(56,8,0)$ $(0,16,48)$ $(24,32,8)$ $(48,16,0)$ $(0,24,40)$ $(20,32,12)$ $(40,24,0)$ $(0,32,32)$ $(16,32,16)$ $(32,32,0)$ $(0,40,24)$ $(12,32,20)$ $(24,40,0)$	$24 = 3\cdot 2^3$

(0, 48, 16) (8, 32, 24) (16, 48, 0)
(0, 56, 8) (4, 32, 28) (8, 56, 0)

Table of Unordered Triplets

Generation #	Unordered triplets	# of unordered triplets
0	{[1, 0, 0]; 3}	1
1	{[4, 0, 0]; 3} {[2, 2, 0]; 2} {[2, 1, 1]; 1}	$3 = 2^1 + 1$
2	{[16, 0, 0]; 3} {[12, 4, 0]; 4} {[8, 8, 0]; 2} {[8, 6, 2]; 2} {[8, 4, 4]; 2}	$5 = 2^2 + 1$
3	{[64, 0, 0]; 3} {[56, 8, 0]; 4} {[48, 16, 0]; 4} {[40, 24, 0]; 4} {[32, 32, 0]; 2} {[32, 28, 4]; 2} {[32, 24, 8]; 2} {[32, 20, 12]; 2} {[32, 16, 16]; 1}	$9 = 2^3 + 1$

The entries in these tables have also been calculated for $n = 4$ with results consistent with the above entries.

4. Observations on the Tables of Triplets.

Four observations can be made about these tables. For values of $n \leq 4$, we have:

1. For generation number $n > 0$, each unordered triplet is a partition into one, two, or three parts of 2^{2n}. For each generation there is a unordered triplet of the form $[2^{2n}, 0, 0]$.

This observation for general n is proved as follows. Make the induction assumption that

$$\delta_n + \mu_n + \rho_n = 2^{2n}.$$

This relation holds for $n = 1$. By use of the multiplication table (1) it can be checked easily that for the $n + 1$ generation the unordered triplet $(\delta_{n+1}, \mu_{n+1}, \rho_{n+1})$ satisfies

$$\delta_{n+1} + \mu_{n+1} + \rho_{n+1} = 4\ 2^{2n} = 2^{2(n+1)}.$$

The existence of the unordered triplet of the form $(2^{2n}, 0, 0)$ follows from the existence of the zygote $d^n = 2^{2n}d$. Thus observation 1 holds for general n.

2. An example of a triplet that, because of our special mating rules, does not appear is (12,40,12) which arises from $(dm)(mr) \neq d(m(m(r)))$.

3. For an unordered triplet of the form $[a,0,0]$, the multiplicity is 3. For an unordered triplet of the form $[a,a,0]$, the multiplicity is 2. For an unordered triplet of the form $[a,b,0]$, $a \neq b$, the multiplicity is 4. For an unordered triplet of the form $[a,b,c]$, a,b,c all different, the multiplicity is 2. For an unordered triplet of the form $[a,b,b]$, the multiplicity is 1. We have not proved this observation for general n.

4. At generation number $n > 0$, there are only the two sets of unordered triplets:

$$\begin{aligned} S_n^1 &: [(2^n - k)2^n, k2^n, 0], \qquad 0 \leq k \leq 2^{n-1} \\ S_n^2 &: [2^{2n-1}, (2^n - k)2^{n-1}, k2^{n-1}], \qquad 1 \leq k \leq 2^{n-1} \end{aligned} \tag{4}$$

This observation is stated in Theorem 2 below for general n.

5. $3 \cdot 2^n$ Ordered Triplets are Present at the nth Generation.

We first examine the types of ordered triplets present. It appears that the triplets present in the above table of ordered triplets are of four types at the nth generation:

Type 1_n. $(2^{2n}, 0, 0), (0, 2^{2n}, 0), (0, 0, 2^{2n})$.

Type 2_n. $((2^n - k)2^n, k2^n, 0)$, $1 \leq k \leq 2^n - 1$.

Type 3_n. $(0, k2^n, (2^n - k)2^n)$, $1 \leq k \leq 2^n - 1$.

Type 4_n. $(k2^{n-1}, 2^{2n-1}, (2^n - k)2^{n-1})$, $1 \leq k \leq 2^n - 1$.

To do the induction proof for Theorem 1, we repeat the above list of types for the $n+1$st generation:

Type 1_{n+1}. $(2^{2(n+1)}, 0, 0), (0, 2^{2(n+1)}, 0), (0, 0, 2^{2(n+1)})$.

Type 2_{n+1}. $((2^{n+1} - k)2^{n+1}, k2^{n+1}, 0)$, $1 \leq k \leq 2^{n+1} - 1$.

Type 3_{n+1}. $(0, k2^{n+1}, (2^{n+1} - k)2^{n+1})$, $1 \leq k \leq 2^{n+1} - 1$.

Type 4_{n+1}. $(k2^n, 2^{2n+1}, (2^{n+1} - k)2^n)$, $1 \leq k \leq 2^{n+1} - 1$.

Theorem 1. The branching process defined by (7) below with the set of ordered triplets at the 0-th generation given as

$$[1,0,0],\ [0,1,0];\ [0,0,1] \tag{5}$$

is the set of ordered triplets at the n-th generation given by

$$(1_n,\ 2_n,\ 3_n, 4_n). \tag{6}$$

These ordered triplets yield distinct ratios. Their cardinality is $3 \cdot 2^n$.

Proof. To show the relations between the two generations, it is useful to spell out the relations between ordered triplets at generation n: $(\delta_n, \mu_n, \rho_n)$ and the ordered triplets at generation $n+1$: $(\delta_{n+1}, \mu_{n+1}, \rho_{n+1})$. These relations are given by:

$$\begin{aligned} d(\delta_n d + \mu_n m + \rho_n r) &= 2(2\delta_n + \mu_n)d + 2(\mu_n + 2\rho_n)m \\ m(\delta_n d + \mu_n m + \rho_n r) &= (2\delta_n + \mu_n)d + 2(\delta_n + \mu_n + \rho_n)m + (\mu_n + 2\rho_n)r \\ r(\delta_n d + \mu_n m + \rho_n r) &= 2(2\delta_n + \mu_n)m + 2(\mu_n + 2\rho_n)r \end{aligned}$$

We note that these relations are invariant under interchange of d and r. The relations can also be written in the form

$$\begin{aligned} d(\delta_n, \mu_n, \rho_n) &= (2(2\delta_n + \mu_n), 2(\mu_n + 2\rho_n), 0) \\ m(\delta_n, \mu_n, \rho_n) &= (2\delta_n + \mu_n, 2(\delta_n + \mu_n + \rho_n), \mu_n + 2\rho_n) \\ r(\delta_n, \mu_n, \rho_n) &= (0, 2(2\delta_n + \mu_n)), 2(\mu_n + 2\rho_n)) \end{aligned} \tag{7}$$

In (7), d, m, r can be regarded as operators that take an ordered triplet in the nth generation into an ordered triplet in the $n+1$th generation. We shall now apply the operators d, m, r to each of the ordered triplets of types 1_n, 2_n, 3_n and 4_n. We obtain the following results. The type of the generated ordered triplet in the $n+1$ generation is given to the right.

For Type 1_n triplets we have:

$$\begin{aligned} d(2^{2n}, 0, 0) &= (2^{2(n+1)}, 0, 0) && 1_{n+1} && (8a) \\ d(0, 2^{2n}, 0) &= (2^{2n+1}, 2^{2n+1}, 0) && 2_{n+1} && (8b) \\ d(0, 0, 2^{2n}) &= (0, 2^{2(n+1)}, 0) && 1_{n+1} && (8c) \\ m(2^{2n}, 0, 0) &= (2^{2n+1}, 2^{2n+1}, 0) && 2_{n+1} && (8d) \\ m(0, 2^{2n}, 0) &= (2^{2n}, 2^{2n+1}, 2^{2n}) && 4_{n+1} && (8e) \\ m(0, 0, 2^{2n}) &= (0, 2^{2n+1}, 2^{2n+1}) && 3_{n+1} && (8f) \\ r(2^{2n}, 0, 0) &= (0, 2^{2(n+1)}, 0) && 1_{n+1} && (8g) \\ r(0, 2^{2n}, 0) &= (0, 2^{2n+1}, 2^{2n+1}) && 3_{n+1} && (8h) \\ r(0, 0, 2^{2n}) &= (0, 0, 2^{2(n+1)}) && 1_{n+1} && (8k) \end{aligned}$$

It is seen that the set of right-hand sides include the set Type 1_{n+1}.

For Type 2_n triplets we have:

$$\begin{aligned} d((2^n - k)2^n, k2^n, 0) =& ((2^{n+1} - k)2^{n+1}, k2^{n+1}, 0), \quad 2_{n+1} \\ & 1 \le k \le 2^n - 1, \end{aligned} \tag{9a}$$

$$m((2^n-k)2^n, k2^n, 0) = (k'2^n, 2^{2n+1}, (2^{n+1}-k')2^n), \quad 4_{n+1}$$
$$k' = 2^{n+1}-k,\ 1 \le k \le 2^n-1,\ 2^n+1 \le k' \le 2^{n+1}-1, \tag{9b}$$

$$r((2^n-k)2^n, k2^n, 0) = (0, k'2^{n+1}, (2^{n+1}-k')2^{n+1}), \quad 3_{n+1}$$
$$k' = 2^{n+1}-k,\ 1 \le k \le 2^n-1,\ 2^n+1 \le k' \le 2^{n+1}-1. \tag{9c}$$

For Type 3_n triplets we have

$$\begin{aligned} d(0, k2^n, (2^n-k)2^n) &= ((2^{n+1}-k')2^{n+1}, k'2^{n+1}, 0), \quad 2_{n+1} \\ k' &= 2^{n+1}-k,\ 1 \le k \le 2^n-1,\ 2^n+1 \le k' \le 2^{n+1}-1, \end{aligned} \tag{9d}$$

$$\begin{aligned} m(0, k2^n, (2^n-k)2^n) =& (k2^n, 2^{2n+1}, (2^{n+1}-k)2^n), \quad 4_{n+1} \\ & 1 \le k \le 2^n-1, \end{aligned} \tag{9e}$$

$$\begin{aligned} r(0, k2^n, (2^n-k)2^n) =& (0, k2^{n+1}, (2^{n+1}-k)2^{n+1}), \quad 3_{n+1} \\ & 1 \le k \le 2^n-1. \end{aligned} \tag{9f}$$

For Type 4_n triplets we have

$$d(k2^{n-1}, 2^{2n-1}, (2^n-k)2^{n-1}) = ((2^{n+1}-k')2^{n+1}, k'2^{n+1}, 0), \quad 2_{n+1},$$
$$k' = 3\cdot 2^{n-1}-k,\ 1 \le k \le 2^n-1,\ 2^{n-1}+1 \le k' \le 3\cdot 2^{n-1}-1, \tag{9g}$$

$$m(k2^{n-1}, 2^{2n-1}, (2^n-k)2^{n-1})) = (k'2^n, 2^{2n+1}, (2^{n+1}-k')2^n), \quad 4_{n+1}$$
$$k' = 3\cdot 2^{n-1}-k,\ 1 \le k \le 2^n-1,\ 2^{n-1}+1 \le k' \le 3\cdot 2^{n-1}-1, \tag{9h}$$

$$r(k2^{n-1}, 2^{2n-1}, (2^n-k)2^{n-1})) = (0, k'2^{n+1}, (2^{n+1}-k')2^{n+1}), \quad 3_{n+1}$$
$$k' = 2^{n-1}+k,\ 1 \le k \le 2^n-1\,;\ 2^{n-1}+1 \le k' \le 3\cdot 2^{n-1}-1. \tag{9k}$$

We see that (9a), (9d), and (9g) give exactly all of the type 2_{n+1} set. We see that (9b), (9e), and (9h) give exactly all of the type 3_{n+1} set. We see that (9c), (9f), and (9k) give exactly all of the type 4_{n+1} set. Thus the set $(1_{n+1},\ 2_{n+1},\ 3_{n+1}, 4_{n+1})$ is the $n+1$ generation, provided the induction hypothesis holds for $n = 0$. It can be checked that the hypothesis does hold for $n = 0$. This completes the proof of the theorem.

It can be checked that the set in (6), when arranged as unordered triplets is exactly the sets S_n^1 and S_n^2 in (4) of which there are $2^n + 1$ members. Thus we have the following theorem:

Theorem 2. The set of unordered triplets arising from the branching process described in Theorem 1 have $2^n + 1$ members at the n-th generation.

6. Discussion

The manipulations in the induction proof of Theorem 1 prove the intuitively obvious conclusion that iteratively filling three ordered slots with choices between two genes, n times, leads to $3 \cdot 2^n$ ordered combinations. Theorem 2 gives the size of the set of unordered triplets at the nth generation.

Acknowledgement

The authors thank Darryl Holm for valuable help and advice on this project.

References

1. Andrews, George *The Theory of Partitions*, Addison-Wesley Publishing Company, 1976.
2. Etherington, I. M. H., Genetic algebras, Proc. Roy. Soc. Edinburgh, **59**, 242-258, 1939.
3. Etherington, I. M. H., Non-Associative Combinations, Proc. Roy. Soc. Edinburgh, **59**, 153-162, 1939.
4. Hille, Einar and Ralph S. Phillips, *Functional Analysis and Semi-groups*, American Mathematical Society, 1957.
5. Jenks, Richard D. and Robert S. Sutor, *Axiom, The Scientific Computation System*, Springer-Verlag, 1992.
6. Wörz-Busekros, Angelika, *Algebras in Genetics*, Springer-Verlag, Lecture Notes in Biomathematics, Vol. 36, 1980.