SOME APPLICATIONS OF MATHEMATICS TO BREEDING PROBLEMS

RAINARD B. ROBBINS
University of Michigan, Ann Arbor, Michigan

[Received March 23, 1917]

TABLE OF CONTENTS

INTRODUCTION

In a recent paper Professor JENNINGS (1916) has given formulae for the calculation of the results of various systems of breeding in which a single Mendelian trait is in question. It seems that JENNINGS's method gave him no absolute assurance of the correctness of his formulae. To quote from his paper (1916, page 62),

"After a law or regular series is obtained that fits the first five or six generations, the law is applied to give the results for three or four generations more. These results are then tested by the actual detailed working out (symbolic formation of gametes and their mating, etc.) for these same later generations; if the formula has given the correct results, it is assumed to be a general formula."

Again (1916, page 61),

"I am compelled, therefore, in most cases, to content myself with giving the actual formulae, leaving their correctness to the test of time."

It is the purpose of this paper, first, to give some examples to show how a method of mathematical repetition can be used to suggest formulae and how mathematical induction can be used to establish a formula when once suggested; second, to express the nth term of series in JENNINGS's table 1, (1916, page 54) as a function of n; third, to solve the problem of inbreeding by brother and sister mating. This paper deals only with a single pair of typical Mendelian factors.

PART I. APPLICATIONS OF THE METHODS OF MATHEMATICAL INDUCTION AND REPETITION

1. *Random mating in a general population*

Consider the problem of random mating in a population consisting of $r\,AA + s\,Aa + t\,aa$. The fundamental method of considering all possible crosses gives the results stated by JENNINGS (1916, page 65) for the first generation:

1) $(s + 2r)^2AA + 2(s + 2r)(s + 2t)Aa + (s + 2t)^2aa.$

It should be stated once for all that it is only the relative magnitudes of the coefficients of AA, Aa and aa which are of interest. It has been shown[1] that 1) gives the result for all following generations. A proof will be given here to illustrate a method which is quite valuable for other problems in breeding.

To get the composition of the second generation, one should note that he has merely a repetition of the problem of getting the composition of the first generation. We have to consider the problem of random mating in a population consisting of $R\,AA + S\,Aa + T\,aa$, in which

2) $R = (s + 2r)^2, S = 2(s + 2r)(s + 2t), T = (s + 2t)^2.$

It is needless to repeat the work involved in obtaining expression 1). We read from 1) immediately that the second generation will have the composition

3) $(S + 2R)^2\,AA + 2(S + 2R)(S + 2T)Aa + (S + 2T)^2aa.$

To find what this means in terms of r, s, t, we substitute the values of R, S, T from 2) into the expression 3). This gives the composition

$$AA = 16(r + s + t)^2(s + 2r)^2.$$
$$Aa = 16(r + s + t)^2 2(s + 2r)(s + 2t).$$
$$aa = 16(r + s + t)^2(s + 2t)^2.$$

For want of a better name this process is called "mathematical repeti-

[1] This has been proved by WENTWORTH and REMICK (1916) who state that JENNINGS also had the result.

tion." Omitting the common factor, $16(r+s+t)^2$, which has nothing to do with the proportions involved, we have the same composition for the second generation that we had for the first.

We can read from this result more than a conclusion regarding the second generation. We can say that random mating in any population of composition 1) results in another generation of the same composition. Thus for our original problem, we have the conclusion that after the first random mating the proportions in the population are fixed and are given by expression 1).

2. *A special case of assortative mating*

This example is to illustrate how mathematical induction can be used to test the accuracy of a formula when once suggested. Consider the problem of assortative mating, dominants with dominants, recessives with recessives. Beginning with a cross between AA and aa, and following this by assortative mating for n generations, JENNINGS (1916, page 66) gives the resultant composition as follows:

4) $(n+1)AA + 2Aa + (n+1)aa.$

If this composition is correct for a particular value of n, and assortative mating occurs in the population it represents, the next generation should show a composition obtained from 4) by replacing n with $n+1$. Conversely, if assortative mating in the population 4) gives a population of composition obtained by replacing n by $n+1$ in 4), and if our original problem gives the distribution 4) for $n=1$, then the formula 4) holds for all values of n. The most elementary methods show that 4) holds for $n=1$. Then to complete the proof it is only necessary to show that assortative mating in a population 4) results in a population of composition obtained by replacing n by $n+1$ in 4); i.e.,

5) $(n+2)AA + 2Aa + (n+2)aa.$

In assortative mating the AA and Aa individuals mate at random while the aa individuals mate with like kind. Out of every $2n+4$ children, $n+3$ will come from dominant parents, the remaining $n+1$ coming from recessive parents. The crosses among the dominants will be in the proportions

$$(n+1)^2 AA \times AA,\ 4(n+1)AA \times Aa,\ 4\,Aa \times Aa.$$

We shall use the notation (a, b, c) to indicate a individuals of type AA, b of type Aa and c of type aa. Then the three crosses noted will produce individuals in the following proportions:

$$\begin{array}{llll} & (\ a, & b, & c) \\ (n+1)^2 AA \times AA = & ((n+1)^2, & 0, & 0). \\ 4(n+1) AA \times Aa = & (2(n+1), & 2(n+1), & 0). \\ 4 \quad Aa \times Aa = & (\quad 1, & 2, & 1). \\ \text{Totals} \quad = & ((n+2)^2, & 2(n+2), & 1). \end{array}$$

Then the $(n+1)$th generation consists of individuals in the following proportions:

$$AA = \frac{(n+2)^2}{(n+3)^2} \cdot \frac{n+3}{2n+4}; \; Aa = \frac{2(n+2)}{(n+3)^2} \cdot \frac{n+3}{2n+4};$$

$$aa = \frac{1}{(n+3)^2} \cdot \frac{n+3}{2n+4} + \frac{n+1}{2n+4} = \frac{(n+2)^2}{(n+3)(2n+4)}$$

Removing the common factor $1/[2(n+3)]$ we have

$$(n+2)AA + 2Aa + (n+2)aa,$$

which is identical with expression 5) as was desired.

3. *Assortative mating in a general population*

As a final example illustrating both methods, consider the more general problem of assortative mating of the population

$$r\,AA + s\,Aa + t\,aa.$$

Detailed examination of the crosses involved gives the result stated by JENNINGS (1916, page 67) for the first generation,

6) $(2r+s)^2 AA + 2s(2r+s)Aa + (s^2 + 4rt + 4st)aa.$

The problem is now really simpler than was the special case considered above. To get the composition of the second generation we need not consider the crosses involved at all. If we set

7) $(2r+s)^2 = R,\ 2s(2r+s) = S,\ s^2 + 4rt + 4st = T,$

expression 6) can be written

$$R\,AA + S\,Aa + T\,aa.$$

We seek the result of assortative mating in this population and it is evident that it is only necessary to write expression 6) with large letters. The second generation has the composition,

8) $(2R+S)^2 AA + 2S(2R+S)Aa + (S^2 + 4RT + 4ST)aa.$

To interpret this we must replace R, S, T by their values in r, s, t from equations 7).

$$(2R+S)^2 = 4(2r+s)^2(2r+2s)^2.$$
$$2S(2R+S) = 4(2r+s)^2 2s(2r+2s).$$
$$S^2 + 4RT + 4\,ST = 4(2r+s)(2\,r+2s)(2s^2 + 4rt + 6st).$$

Omitting the common factor $4(2r + s)(2r + 2s)$, we have for the second generation

9) $(2r + s)(2r + 2s)AA + 2s(2r + s)Aa + (2s^2 + 4rt + 6st)aa.$

This, or at least one more repetition of the process, suggests that the nth generation will have the composition[2]

10) $(2r + s)(2r + ns)AA + 2s(2r + s)Aa + [ns^2 + 4rt + 2(n + 1)st]aa.$

Inspection shows that this formula holds for $n = 1$ and $n = 2$. If we assume 10) thinking of n as fixed, and show that assortative mating in such a population gives a generation whose composition is obtained by replacing n by $n + 1$ in 10), then we shall know that 10) holds for all values of n. To do this let

$R = (2r + s)(2r + ns)$; $S = 2s(2r + s)$; $T = ns^2 + 4rt + 2(n + 1)st,$

and form expression 6) in the large letters; i.e., the expression 8) with our present meaning for R, S, T. This process gives for the proportions in the $(n + 1)$th generation.

$$AA = 4(2r + s)^2[2r + (n + 1)s]^2.$$
$$Aa = 4(2r + s)^2.\ 2s[2r + (n + 1)s].$$
$$aa = 4(2r + s)[2r + (n + 1)s][n + 1)s^2 + 4rt + 2(n + 2)st].$$

Dividing by the common factor $4(2r + s)[2r + (n + 1)s]$ the proportions become,

$$(2r + s)[2r + (n + 1)s]AA + 2s(2r + s)Aa + [(n + 1)s^2 + 4rt + 2(n + 2)st]aa.$$

Inspection shows that these results may be obtained by replacing n by $n + 1$ in expression 10).

It should be of interest to note that as n increases indefinitely the proportions in 10) approach the proportions in

$$(2r + s)AA + 0\,Aa + (2t + s)aa.$$

These examples should show, first that the method of mathematical repetition can be used to simplify the work of calculating the composition of higher generations; second, that the method of mathematical induction can be used to prove or disprove a general formula for the composition of the nth generation when it has once been suggested.

PART II. GENERAL TERMS OF JENNINGS'S SERIES

In table 1 JENNINGS (1916, page 54) gives twenty terms of each of several series which present themselves in breeding problems. For series

[2] This result was obtained by WENTWORTH and REMICK (1916).

B, C, D and E he gives the nth term as a function of n. It may be desirable to have the nth term of his other series (lettered from F to M). Inspection shows that only two of these are independent and if we can express the nth term of each of them, the others come immediately. The derivation of these two nth terms will be given next and then the nth term of each series will be written down.

1. *Derivation of the nth term of the Fibonacci series*

The Fibonacci series F is defined by its first two terms, $F_0 = 0$, $F_1 = 1$, and the recurrence relation $F_n = F_{n-1} + F_{n-2}$. In mathematical language we have to solve the homogeneous recurrence equation

11) $F_n - F_{n-1} - F_{n-2} = 0$

with the initial conditions, $F_0 = 0$ and $F_1 = 1$. It is well known that C^n is a solution of 11), where C is a root of $C^2 - C - 1 = 0$; i.e., $C = (1 \pm \sqrt{5})/2$. Then $[(1 + \sqrt{5})/2]^n$ and $[(1 - \sqrt{5})/2]^n$ are solutions of 11) and any solution can be put in the form

12) $F_n = [K_1(1 + \sqrt{5})^n + K_2(1 - \sqrt{5})^n]/2^n$.

We wish to determine the constants K_1 and K_2 so that $F_0 = 0$ and $F_1 = 1$. Setting $n = 0$ and $n = 1$ in equation 12), we have

13) $F_0 = K_1 + K_2 = 0$.

14) $F_1 = [K_1(1 + \sqrt{5}) + K_2(1 - \sqrt{5})]/2 = 1$.

From 13), $K_1 = - K_2$. Substituting in 14),

$$F_1 = K_1[1 + \sqrt{5} - (1 - \sqrt{5})]/2 = 1.$$

$$K_1 = 1/\sqrt{5};\ K_2 = - 1/\sqrt{5};\ \text{and}$$

15) $F_n = [(1 + \sqrt{5})^n - (1 - \sqrt{5})^n]/[\sqrt{5} \cdot 2^n]$.

The rather complicated appearance of this formula may make it seem useless. If one desires only a few of the early terms in the series, it would most certainly not be advisable to use this formula. But suppose you want the 100th term. By using logarithms it is about as easy to get the 100th term with all desirable accuracy from this formula 15) as it is to get the tenth term, and no time need be spent calculating the first 99 terms.

The formula 15) for the Fibonacci series enables us to prove the following important

THEOREM: *As n increases indefinitely, the nth term of the Fibonacci series divided by 2^n approaches zero as a limit.*

Symbolically stated, the theorem is

16) $\lim_{n = \infty} F_n/2^n = 0$.

Writing in the value of F_n this becomes

$$\underset{n=\infty}{\text{Limit}} \frac{(1+\sqrt{5})^n-(1-\sqrt{5})^n}{\sqrt{5}\cdot 4^n}=0.$$

The proof consists in noting that $(1+\sqrt{5})/4$ and $(1-\sqrt{5})/4$ are proper fractions and that as a proper fraction is raised to higher and higher powers, the result approaches zero as a limit. As an immediate corollary we have that if C_1 and c_2 are constants,

$$\underset{n=\infty}{\text{Limit}} \quad C_1F_n/2^{n+c_2}=0.$$

This follows because $C_1/2^{c_2}$ is a constant, say C_3, and we have

$$\underset{n=\infty}{\text{Limit}} \frac{C_3\,F_n}{2^n}=C_3 \underset{n=\infty}{\text{Limit}} \frac{F_n}{2^n}=0.$$

2. *Derivation of series* G

The second series which it is necessary to consider is defined by the recurrence $G_n = 2^{n-1} - G_{n-1}$, together with the initial condition $G_0 = 0$. We have to solve the non-homogeneous recurrence

17) $G_n + G_{n-1} = 2^{n-1}$

subject to the condition $G_0 = 0$. The most general solution is the sum of the general solution of the homogeneous equation

18) $G_n + G_{n-1} = 0$

and any particular solution of equation 17). The general solution of 18) is $K\,C^n$ where C is a solution of $C+1=0$; i.e., $K(-1)^n$. A particular solution of equation 17) is $G_n = 2^n/3$.

The general solution of 17) is therefore

19) $G_n = K(-1)^n + 2^n/3$.

We wish to determine K so that $G_0 = 0$. Setting $n = 0$ in 19), we have

$G_0 = K + 1/3. \therefore K = -1/3$ and

20) $G_n = [2^n - (-1)^n]/3$.

The value of a formula for G_n is particularly apparent in an example given by JENNINGS (1916, page 80). The series $G_n . G_{n+1}/2^{2n-1}$ is needed. Substituting the value of G_n and G_{n+1} this fraction is

$$\frac{1}{9}\left[\frac{2^{2n+1}-(-2^n)-1}{2^{2n-1}}\right]=\frac{4}{9}-\frac{(-2)^n+1}{9\cdot 2^{2n-1}}.$$

From this expression, the various terms of the series can be calculated readily, independently, and without recourse to any complicated rule, and the limit approached as n increases indefinitely is apparent.

3. *The nth terms of series in* Jennings's *table*

Using the values of F_n and G_n we can write down the following set of nth terms for Jennings's series:

$$F_n = \frac{(1+\sqrt{5})^n - (1-\sqrt{5})^n}{\sqrt{5}\,.\,2^n};\ G_n = \frac{1}{3}\left[2^n - (-1)^n\right];\ B_n = 2^n.$$

$$H_n = G_n - F_n = \frac{1}{3}\left[2^n - (-1)^n\right] + \frac{(1-\sqrt{5})^n - (1+\sqrt{5})^n}{\sqrt{5}\,.\,2^n}.$$

$$I_n = B_n - G_n - F_n = \frac{2^{n+1} + (-1)^n}{3} + \frac{(1-\sqrt{5})^n - (1+\sqrt{5})^n}{\sqrt{5}\,.\,2^n}.$$

$$J_n = B_n - F_{n+1} = 2^n + \frac{(1-\sqrt{5})^{n+1} - (1+\sqrt{5})^{n+1}}{\sqrt{5}\,.\,2^{n+1}}.$$

$$K_n = B_n - F_{n+2} = 2^n + \frac{(1-\sqrt{5})^{n+2} - (1+\sqrt{5})^{n+2}}{\sqrt{5}\,.\,2^{n+2}}.$$

$$L^3{}_n = B_n - G_{n-1} - F_{n-1} = \frac{5.\,2^{n-1} + (-1)^{n-1}}{3} + \frac{(1-\sqrt{5})^{n-1} - (1+\sqrt{5})^{n-1}}{\sqrt{5}\,.\,2^{n-1}}.$$

$$M_n = 3\ B_n - F_{n+2} = 3\quad 2^n + \frac{(1-\sqrt{5})^{n+2} - (+\sqrt{5})^{n+2}}{\sqrt{5}\,.\,2^{n+2}}.$$

Incidentally it may be noted that E_n, given by Jennings as $2^{n-1} + 2^{n-2} - 1$ can be written in the slightly more compact form, $E_n = 3.\ 2^{n-2} - 1$.

PART III. BROTHER AND SISTER MATING

1. *Results in random brother and sister mating*

Given a family consisting of $r\,AA + s\,Aa + t\,aa$, what is the composition of the nth generation if mating is restricted to random mating between brothers and sisters? Special cases of this problem have been considered by Jennings (1916) and Pearl (1914).

For the benefit of those who do not care to follow the details of the development, the results will be stated first. The nth generation, i.e., the generation resulting from the nth brother and sister mating, has the following composition:

$AA = [1 + K_2 - T_n]/2$; $Aa = T_n$; $aa = [1 - K_2 - T_n]/2$, in which $T_n = [s\,K\,F_{n+1} + (rs + st + 4rt)F_n]/K^2.2^n$; $K_2 = (r - t)/K$; $K = r + s + t$; and F_n is the general term of the Fibonacci series.

[3] By what is evidently a slip, Jennings writes F_{n+1} in this equation for F_{n-1}.

2. *Development of above results*

Three types of individuals are involved, *AA, Aa* and *aa*. The different possible crosses of individuals of these types together with the composition of the resulting families are given below. The notation (a, b, c) means that individuals of the types *AA, Aa, aa* appear in numbers proportional to a, b, c.

Kind of cross	Composition of resulting family	Letter indicating type of family
$AA \times AA$	(1, 0, 0)	o
$AA \times Aa$	(½, ½, 0)	p
$AA \times aa$	(0, 1, 0)	q
$Aa \times Aa$	(¼, ½, ¼)	r
$Aa \times aa$	(0, ½, ½)	u
$aa \times aa$	(0, 0, 1)	v

It is useful to keep track of these six kinds of families. Let o_n, p_n, q_n, r_n, u_n, v_n be the relative numbers of families of the various kinds in the order given above. If we can calculate $o_n \ldots\ldots v_n$ we can readily find the numbers of *AA, Aa, aa* individuals in the nth generation.

a. *Development of the formulae for* $o_n, p_n, q_n, r_n, u_n, v_n$

To find o_n, for instance, we examine the source of the families in the nth generation of the type o. All the children of families of type o in the $(n—1)$th generation will be in families of type o, since *AA* individuals only are concerned. One-fourth of the families which consist of children of families of type p in the $(n—1)$th generation will be of type o and 1/16 of the families which are children of families of type r in the $(n—1)$th generation will be of type o. Thus we have that[4]

21) $o_n = o_{n-1} + p_{n-1}/4 + r_{n-1}/16.$

Similar considerations give

22) $p_n = p_{n-1}/2 + r_{n-1}/4.$

23) $q_n = r_{n-1}/8.$

24) $r_n = p_{n-1}/4 + q_{n-1} + r_{n-1}/4 + u_{n-1}/4.$

25) $u_n = u_{n-1}/2 + r_{n-1}/4.$

26) $v_n = v_{n-1} + u_{n-1}/4 + r_{n-1}/16.$

The problem before us is to solve this system of recurrence relations.

[4] PEARL (1914) had these equations, except that in the case he considered, $o_n = v_n$; $p_n = u_n$. The notation here used was used by PEARL.

We first set

27) $p_n - u_n = y_n,$ and

28) $o_n - v_n = x_n.$

Then from equations 22) and 25),

29) $y_n = y_{n-1}/2,$ and similarly

30) $x_n = x_{n-1} + y_{n-1}/4$ and the above system, 21)-26), may be replaced by the following system:

21′) $o_n = o_{n-1} + p_{n-1}/4 + r_{n-1}/16.$

22′) $p_n = p_{n-1}/2 + r_{n-1}/4.$

23′) $r_n = r_{n-1}/4 + r_{n-2}/8 + p_{n-1}/2 - y_{n-1}/4.$

24′) $y_n = y_{n-1}/2.$

25′) $x_n = x_{n-1} + y_{n-1}/4.$

Equation 24′) may be written

$$2y_n - y_{n-1} = 0.$$

The most general solution of this equation is

31) $y_n = K_1/2^n$, in which K_1 is an arbitrary constant. Then equation 25′) becomes

$$x_n - x_{n-1} = K_1/2^{n+1}.$$

The most general solution of this equation is

32) $x_n = K_2 - K_1/2^{n+1}$, K_2 being an arbitrary constant.

From equation 22′)

$$33)\quad \begin{cases} r_{n-1} = 4p_n - 2p_{n-1}, \\ r_{n-2} = 4p_{n-1} - 2p_{n-2}, \\ r_n \;\; = 4p_{n+1} - 2p_n. \end{cases}$$

Substituting these values of $r_n, r_{n-1}, r_{n-2},$ in equation 23′) and using equation 31) gives the equation

34) $16p_{n+1} - 12p_n - 2p_{n-1} + p_{n-2} = -K_1/2^{n-1}.$

The corresponding algebraic equation is $16c^3 - 12c^2 - 2c + 1 = 0$; the roots are $c = 1/4$; $c = (1 + \sqrt{5})/4$; $c = (1 - \sqrt{5})/4$. Then the most general solution of the homogeneous equation

$$16p_{n+1} - 12p_n - 2p_{n-1} + p_{n-2} = 0 \text{ is}$$

$$[K_3(1 + \sqrt{5})^n + K_4(1 - \sqrt{5})^n + K_5]/4^n,$$

in which K_3, K_4 and K_5 are arbitrary constants. A particular solution of the non-homogeneous equation 34) is $K_1/2^{n+1}$. Therefore the general solution of equation 34) is

$$35)\quad p_n = \frac{K_1}{2^{n+1}} + \frac{K_3(1 + \sqrt{5})^n + K_4(1 - \sqrt{5})^n + K_5}{4^n}$$

Let $P_n = K_3(1 + \sqrt{5})^n + K_4(1 - \sqrt{5})^n$.
Then 35) may be written,

36) $p_n = K_1/2^{n+1} + (P_n + K_5)/4^n$.

From $y_n = p_n - u_n$, we have $u_n = p_n - y_n = p_n - K_1/2^n$.

37) $u_n = -K_1/2^{n+1} + (P_n + K_5)/4^n$.

From 33) $r_n = 4p_{n+1} - 2p_n$. A little algebraic reduction shows that this becomes

38) $r_n = [4P_{n-1} - K_5]/4^n$.

Since $q_n = r_{n-1}/8$, we have

39) $q_n = (4P_{n-2} - K_5)/2 \times 4^n$.

By direct substitution one can verify that

$$P_n - 2P_{n-1} - 4P_{n-2} = 0.$$

Using this equation, q_n may be written

40) $q_n = [P_n - 2P_{n-1} - K_5]/2 \times 4^n$.

Finally, to get o_n and v_n we note that since $o_n \ldots\ldots v_n$ are only proportional to the numbers of families of different types, it will simplify the problem to choose them so that,

$$o_n + p_n + q_n + r_n + u_n + v_n = 1.$$

Then $o_n + v_n = 1 - (p_n + q_n + r_n + u_n)$.

From equation 32) we have that

$$o_n - v_n = K_2 - K_1/2^{n+1}.$$

Solving the last two equations for o_n and v_n,

41) $o_n = -K_1/2^{n+2} + \dfrac{K_2 + 1}{2} - \dfrac{1}{2}(p_n + q_n + r_n + u_n)$.

42) $v_n = K_1/2^{n+2} - \dfrac{K_2 - 1}{2} - \dfrac{1}{2}(p_n + q_n + r_n + u_n)$.

Substituting the values of p_n, q_n, r_n, u_n, from equations 36), 37), 38), 40) into equations 41), 42) gives,

43) $o_n = \dfrac{1 + K_2}{2} - \dfrac{K_1}{2^{n+2}} - \dfrac{5P_n + 6P_{n-1} + K_5}{4^{n+1}}$.

44) $v_n = \dfrac{1 - K_2}{2} + \dfrac{K_1}{2^{n+2}} - \dfrac{5P_n + 6P_{n-1} + K_5}{4^{n+1}}$.

The constants $K_1 \ldots\ldots K_5$ are to be found in terms of the initial conditions; in our problem they are functions of r, s, t. To determine them we need the values of o_1, p_1, q_1, r_1, u_1, v_1. Considering the possible crosses involved in mating the family $r\,AA + s\,Aa + t\,aa$ and using the notation $K = r + s + t$, we find that

$$o_1 = \frac{r^2}{K^2}; \; p_1 = \frac{2rs}{K^2}; \; q_1 = \frac{2rt}{K^2};$$
$$r_1 = \frac{s^2}{K^2}; \; u_1 = \frac{2st}{K^2}; \; v_1 = \frac{t^2}{K^2}.$$

To evaluate K_1 we note from equation 31) that $y_1 = K_1/2$. Also $y_1 = p_1 - u_1$ by definition. Then $K_1 = 2y_1 = 2(p_1 - u_1)$ and substituting for p_1, u_1,

45) $$K_1 = \frac{4\,s(r-t)}{K^2}.$$

From equation 32), $K_2 = x_1 + K_1/4 = o_1 - v_1 + K_1/4$; and substituting for o_1, v_1,

46) $$K_2 = \frac{r-t}{K}.$$

More complicated work of the same nature gives for the remaining constants,

47) $$K_3 = \frac{(1+\sqrt{5})s}{5K} + \frac{(1-\sqrt{5})}{5K^2}(s^2 - 4rt).$$

48) $$K_4 = \frac{(1-\sqrt{5})s}{5K} + \frac{(1+\sqrt{5})}{5K^2}(s^2 - 4rt).$$

49) $$K_5 = \frac{4}{5K^2}[s(2r - s + 2t) - 8rt].$$

It should be noted that we have here five constants $K_1 \ldots . K_5$ expressed in terms of three initial numbers r, s, t. This indicates that our method is useful for a more general problem than the one to which it is here applied. This is shown clearly by expressing $K_1 \ldots . K_5$ in terms of o_1, $p_1 \ldots . . v_1$ as follows:

$$K_1 = 2(p_1 - u_1); \; K_2 = o_1 - v_1 + (p_1 - u_1)/2.$$
$$K_3 = [(1+\sqrt{5})(p_1 + u_1) + 4(\sqrt{5} - 1)q_1 + 4r_1]/10.$$
$$K_4 = [(1-\sqrt{5})(p_1 + u_1) - 4(\sqrt{5} + 1)q_1 + 4r_1]/10.$$
$$K_5 = 4[p_1 + u_1 - 4q_1 - r_1]/5.$$

With this set of values of $K_1 \ldots . K_5$ our formulae will give the composition of the population after $n - 1$ brother and sister matings starting with families of the six special types in numbers proportional to o_1, p_1, q_1, r_1, u_1, v_1.

b. Proportions of the three types of individuals in the nth generation

The final results desired are the numbers giving the proportions of

AA, Aa, aa individuals in the nth generation. It is readily seen that they are[5]

$$AA = o_n + p_n/2 + r_n/4.$$
$$Aa = (p_n + r_n + 2q_n + u_n)/2.$$
$$aa = r_n/4 + u_n/2 + v_n.$$

Substituting the values of $o_n \ldots v_n$,

50) $$AA = \frac{1 + K_2}{2} - \frac{3P_n + 2P_{n-1}}{4^{n+1}}.$$

51) $$Aa = \frac{3P_n + 2P_{n-1}}{2 \times 4^n}.$$

52) $$aa = \frac{1 - K_2}{2} - \frac{3P_n + 2P_{n-1}}{4^{n+1}}.$$

The expression $3P_n + 2P_{n-1}$ which enters these three equations is

$$3P_n + 2P_{n-1} = \frac{\sqrt{5}}{2}[K_3(1 + \sqrt{5})^{n+1} - K_4(1 - \sqrt{5})^{n+1}].$$

It is instructive to get the proportions in 50), 51), 52) in another form by substituting the values of K_3 and K_4 from equations 47) and 48). This gives

$$Aa = \frac{1}{2^n}[\frac{s\,F_{n+2}}{K} - \frac{s^2 - 4rt}{K^2}F_n].$$

in which F_n is the nth term of the Fibonacci series. Since $F_{n+2} = F_{n+1} + F_n$,

53) $Aa = [sK\,F_{n+1} + (rs + st + 4rt)F_n]/[2^n.\,K^2]$.

in which $K = r + s + t$.

From this form we can read the following results:

1. *If the numbers representing the proportions of* Aa *individuals in successive generations be written with* 2^n *in the denominators, the numerators will satisfy the recurrence,*

$$N_n = N_{n-1} + N_{n-2}.$$

2. *If* $s = 0$ *or* $s^2 = 4rt$, *and the denominators are chosen as* $2^nK^2/(s^2 - 4rt)$ *or* $2^nK/s$, *the numerators will be terms of the Fibonacci series.*

3. *As the number of generations increases, the proportion of heterozygous individuals approaches zero regardless of the values of* r, s, t.

4. *As the number of generations increases, the ratio of* AA *to* aa *indi-*

[5] PEARL (1914) had this result for AA but seems to have erred in getting the numbers for Aa. In the case he considered, $o_n = v_n$ and $p_n = u_n$.

viduals approaches $(2r + s)/(2t + s)$, which is the same as the ratio of A and a gametes in the original family.

c. *Illustrative example*

As a check on these formulae, and to illustrate their application, let us take a special case considered by JENNINGS (1916). Let AA and aa be crossed and assume brother and sister mating thereafter. The children of the original cross are all of type Aa. It is with crosses of these individuals that our problem begins. We therefore have $r = t = 0; s = 1$. Substituting in equations 45) — 49),

$K_1 = K_2 = 0; K_3 = K_4 = 2/5; K_5 = -4/5.$

Substituting these values of the constants into equations 50), 51), 52), and using the notation of part II, $F_n = [(1 + \sqrt{5})^n - (1 - \sqrt{5})^n] / \sqrt{5} \cdot 2^n$,

$$AA = \tfrac{1}{2} - F_{n+1}/2^{n+1};\ Aa = F_{n+1}/2^n;$$
$$aa = \tfrac{1}{2} - F_{n+1}/2^{n+1}.$$

These results agree with JENNINGS's series.

3. *Assortative brother and sister mating*

Given a family consisting of $r\,AA + s\,Aa + t\,aa$, what is the composition of the nth generation if mating is restricted (1) to brothers with sisters and (2) to dominants with dominants and recessives with recessives?

To derive the recurrence relations upon which the solution of this problem depends we note:

a) Families of type q will not appear since they arise only by a cross between AA and aa.

b) Families of type u will not appear since they arise only by a cross between Aa and aa.

c) Random mating will occur in families of types o, p, v.

d) Assortative mating will occur in families of type r, ¾ of the resulting families being of type o, p, r, in the proportion 1 : 4 : 4 and ¼ being of the type v.

These considerations lead to the following equations:

$$54) \quad o_n = o_{n-1} + \frac{p_{n-1}}{4} + \frac{r_{n-1}}{12}.$$

$$55) \quad p_n = \frac{p_{n-1}}{2} + \frac{r_{n-1}}{3}.$$

56) $r_n = \frac{p_{n-1}}{4} + \frac{r_{n-1}}{3}$.

57) $v_n = v_{n-1} + \frac{r_{n-1}}{4}$.

The problem of solving this system of equations is very similar to the problem considered above in studying random brother and sister mating. Using the notation

$$P_n = K_1(5+\sqrt{13})^n + K_2(5-\sqrt{13})^n,$$

the solution takes the form,

58) $o_n = 1 - K_3 - 3\,P_{n+1}/2 \times 12^{n+1}$.

59) $p_n = P_n/12^n$.

60) $r_n = (P_{n+1} - 6P_n)/4 \times 12^n$.

61) $v_n = K_3 - (P_{n+1} - 4P_n)/8 \times 12^n$.

The proportions of the three types of individuals in the nth generation are given by

62) $AA = o_n + \frac{p_n}{2} + \frac{r_n}{4} = 1 - K_3 - \frac{P_{n+1} - 2P_n}{16 \,.\, 12^n}$.

63) $Aa = \frac{p_n + r_n}{2} = \frac{P_{n+1} - 2P_n}{8 \,.\, 12^n}$.

64) $aa = v_n + \frac{r_n}{4} = K_3 - \frac{P_{n+1} - 2P_n}{16 \,.\, 12^n}$.

We have to determine the constants K_1, K_2, K_3 in terms of the initial numbers r, s, t. First, substituting $n = 1$ in equations 58), 59), 60), 61), and solving,

65) $K_1 = 2[(\sqrt{13} - 2)p_1 + (5 - \sqrt{13})r_1]/\sqrt{13}$.

66) $K_2 = 2[(\sqrt{13} + 2)p_1 - (5 + \sqrt{13})r_1]/\sqrt{13}$.

67) $K_3 = [2(1 + v_1 - o_1) - p_1]/4$.

Examination of the first matings shows that

68) $o_1 = \frac{r^2}{(r+s)K}$; $p_1 = \frac{2rs}{(r+s)K}$; $r_1 = \frac{s^2}{(r+s)K}$; $v_1 = \frac{t}{K}$,

in which $K = r + s + t$.

Substituting these values into equations 65), 66), 67), we have,

69) $K_1 = 2s[2r(\sqrt{13} - 2) + s(5 - \sqrt{13})]/[\sqrt{13} \,.\, K(r+s)]$.

70) $K_2 = 2s[2r(\sqrt{13} + 2) - s(\sqrt{13} + 5)]/[\sqrt{13} \,.\, K(r+s)]$.

71) $K_3 = (2t + s)/2K$.

The expressions for AA, Aa, aa, in terms of r, s, t, are far from neat.

The one for Aa will be given; those for AA and aa can be readily calculated from the one for Aa by using equations 62), 63), 64).

$$72)\quad Aa = \frac{s}{2K\sqrt{13}\,.\,12^n\,(r+s)}\Big[(5+\sqrt{13})^n\{(7+\sqrt{13})r+(1+\sqrt{13})s\}+ (5-\sqrt{13})^n\{(-7+\sqrt{13})r-(1-\sqrt{13})s\}\Big].$$

It is instructive to note that

$$\underset{n=}{\text{Limit}}\ \frac{P_n}{12^{n+c}} = 0.$$

This follows from the fact that $(5+\sqrt{13})/12$ and $(5-\sqrt{13})/12$ are proper fractions, and that a proper fraction raised to higher and higher powers approaches zero as a limit. With this in mind we see at once from equation 63) that the proportion of heterozygotes approaches zero as n increases. Then

$$\underset{n=\infty}{\text{Limit}}\ (AA)_n = 1 - K_3 = \frac{2r+s}{2K}, \text{ and}$$

$$\underset{n=\infty}{\text{Limit}}\ (aa)_n = K_3 = \frac{2t+s}{2K}.$$

Here again we see what has been true of every problem in inbreeding, that the heterozygotes tend to disappear and the homozygotes approach the proportion

$$AA/aa = (2r+s)/(2t+s).$$

This is to be expected. In fact the following statement of the case seems obvious:

Any method of breeding which gives A and a gametes equal chances of mating and which tends to eliminate heterozygous individuals will in successive generations give populations which approach a stable condition in which the two types of homozygous individuals appear in the same proportion as were their types of gametes in the original population.

LITERATURE CITED

JENNINGS, H. S., 1916 The numerical results of diverse systems of breeeding. Genetics **1**: 53-89.

PEARL, RAYMOND, 1914 On a general formula for the constitution of the *n*th generation of a Mendelian population in which all matings are of brother × sister. Amer. Nat. **48**: 491-494.

WENTWORTH, E. N., and REMICK, B. L., 1916 Some breeding properties of the generalized Mendelian population. Genetics **1**: 608-616.