June 17, 1940

Dear Crosby,

On looking at your selection problem with Primula it struck me you might be glad to have a note of a process rather generally applicable for solving such problems numerically. I can illustrate it with the case of complete lethality of the homostyle factor giving equal production of viable seed to all genotypes. I think you said you had not studied this case.

The essence of the method is to evaluate rather carefully the path by which the population approaches or leaves its point of equilibrium. If this is done, it should be free from the initial disturbance which you noticed at the beginning of your runs of trial computations, and which must be due to the choice of initial constitution for the population, which is in fact not one of the constitutions through which the species would pass.

For example, in our case we have the equations

\[ P' = \frac{1}{2} - \frac{1}{6} R \]
\[ Q' = \frac{Q}{2(1-P)} \]
\[ R' = \frac{1}{6} R + 2 \frac{R}{2(1-P)} \]

where P, Q and R are the relative frequencies in one generation, and
P', Q' and R' are those of the next.

To see how the thing starts, we have the initial equilibrium values \( P = \frac{1}{2}, \ Q = \frac{1}{2}, \ R = 0 \). At the limit the third equation gives
\[
R' = \frac{7}{6} R
\]
so we put \( x = (\frac{7}{6})^n \) and seek any expansion for \( P \) and \( R \) which satisfies the equation and our power series in \( x \).

In general the value we have found to be \( \frac{7}{6} \) is the largest root of the equation from which it is obtained, and of course difficulties are encountered if one of the other roots is chosen.

Probably you will follow the procedure most easily if I set out the numerical values I have obtained - but not checked - for the coefficients of these expansions, which are obtained one by one by a cyclic iterative process involving values previously obtained. As in this process I have used the corresponding expansion of \( \frac{1}{2(1-F)} \), I give the coefficients of this in a parallel column, so that you can, if you like, verify each step in turn:

<table>
<thead>
<tr>
<th>( P )</th>
<th>( R )</th>
<th>( \frac{1}{2(1-F)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x^0 )</td>
<td>.5</td>
<td>-</td>
</tr>
<tr>
<td>( x^1 )</td>
<td>- .16</td>
<td>1.16</td>
</tr>
<tr>
<td>( x^2 )</td>
<td>24489, 79592</td>
<td>-2</td>
</tr>
<tr>
<td>( x^3 )</td>
<td>34073, 71471</td>
<td>3.24646, 7818</td>
</tr>
<tr>
<td>( x^4 )</td>
<td>45943, 91276</td>
<td>-5.10700, 6228</td>
</tr>
<tr>
<td>( x^5 )</td>
<td>65769, 82009</td>
<td>7.87157, 1601</td>
</tr>
<tr>
<td>( x^6 )</td>
<td>.79634, 77094</td>
<td>12.04854, 831</td>
</tr>
</tbody>
</table>

Given the coefficient \( R \) on any line, designated by \( r_t \), the coefficient of \( x^t \) in the expansion, it follows from the first equation
that
\[ \frac{7^t p_t}{6^{t-1}} = -\frac{1}{6} r_t \]
or
\[ \frac{p_t}{6^{t-1}} = \frac{r_t}{-7^t} \quad \frac{q_t}{7^t-6^{t-1}} \]

At the start, since \( n \), designating the generation number, is measured from an arbitrary origin, we can give what value we like to the three equal fractions
\[ \frac{p_1}{1}, \quad \frac{r_1}{-7}, \quad \frac{q_1}{6} \]

and I have taken these to be each equal to \(-\frac{1}{6}\), since with this choice the subsequent coefficients do not change very violently in value. There is, of course, no need to evaluate all three series of coefficients, since any two of them determine the third.

After obtaining \( p_t \), it is used at once to give the corresponding coefficient, say \( a_t \), in the expansion of \( \frac{1}{2(1-p)} \). Then \( r_{t+1} \) is obtained from the third equation in the form
\[ r_{t+1} = \frac{6^{t+1}}{7(7^t-6^t)} \left( r_{1a_t} + r_{2a_t-1} + \ldots + r_{t}a_1 \right) \]
so starting the next cycle of calculations.

Something like six terms of such expansions are useful, and, though there may well be cases in which more necessary, you will see that the expansions appear to be sufficiently convergent if \( x \) is as small as .1, a value corresponding to which it is easy to calculate the generation number. Having a well determined value on the right path one can fairly rapidly compute a number of successive generations, sufficient to reach the region of effective convergence of the corresponding formula for the approach to the position of final equilibrium.
In this case \( P = \frac{2}{5}, Q = 0, R = \frac{3}{5} \), in the neighbourhood of which \( x = (\frac{6}{5})^{-n'} \), where \( n' \) differs from \( n \) by a constant which can be determined by stepping across from one region of convergence to the other. As the selective intensities of this case are only about one half those of the case I talked about at the Genetical Society, it would not be surprising if about 35 steps were needed; but the recurrence relationship is such that these steps are really very easy. The case in which the heterozygous homostyle bears only three quarters the number of viable seeds borne by other types may be a little more difficult in this respect owing to the general factor \( 1 - \frac{1}{9} R \).

Let me know if you try this, as in my experience problems of this type seem to yield to reason rather nicely when treated in this way.

Yours sincerely,