A general theory of the distribution of gene frequencies
I. Overlapping generations

BY P. A. P. MORAN
Australian National University, Canberra

(Communicated by M. L. Oliphant, F.R.S.—Received 14 August 1957)

The distribution of gene frequency at a single locus in a population of diploid individuals, with two sexes, subject to mutation, non-random mating and phenotypic selection, is obtained in the case where the generations are overlapping so that individuals die one by one. This distribution is of the same form as that obtained by heuristic methods by S. Wright in a randomly mating population but the coefficients are altered both by the non-randomness of the mating and the overlapping of the generations.

In this paper, which is a sequel to two other papers (Moran 1958a, b) a general theory of the distribution of gene frequencies at a single locus is considered which takes account of mutation, non-random mating, phenotypic selection and the fact that the individuals concerned are both diploid and of two sexes. This extends the previous discussion (Moran 1958b) which ignored the possible existence of selection and non-random mating. In a second paper the same problem will be considered in the case where all the individuals of each generation are produced simultaneously. This is the case considered in Wright’s papers (for references see Wright 1939).

The total population size, \( N \), is taken as constant and equal to \( N_1 + N_2 \), where \( N_1 \) and \( N_2 \) are the constant numbers of males and females. The individuals in the population die one by one and are replaced by new individuals formed by the union of gametes, one of which is chosen from the gametic output of the previously existing male population and one similarly from the female population.

We consider a single locus at which there is one of two alleles and thus the \( NV \) diploid individuals in the population have one of the genotypes \( aa \), \( Aa \) or \( AA \). We suppose that the numbers of these among the male population are \( k \), \( N_1 - k - l \) and \( l \), respectively, and similarly \( r \), \( N_2 - r - s \) and \( s \) amongst the females. At gametogenesis we suppose that a gamete \( a \) has a probability \( \alpha_1 \) of mutating to \( A \), and \( A \) similarly has a probability \( \alpha_2 \) of mutating to \( a \). In order to obtain a non-trivial distribution as the population size increases we shall have to suppose that \( \alpha_1 \) and \( \alpha_2 \) become small and we shall write \( \alpha_1 = \beta_1 N^{-1}, \alpha_2 = \beta_2 N^{-1}, \) where \( \beta_1 \) and \( \beta_2 \) are fixed positive quantities.

We must now introduce selection and to do this we first suppose the process is one with a continuous time parameter. We suppose that in any interval of time \( (t, t + dt) \) any individual has the probability \( \lambda_i dt + o(dt) \) of dying where \( i = 1, 2, 3 \) according as the individual has genotype \( aa \), \( Aa \) or \( AA \). Then each individual has a life-time with a negative exponential distribution and an expectation of \( \lambda_i^{-1} \), \( \lambda_2^{-1} \) or \( \lambda_3^{-1} \) according to its genotype. Starting from any point on the time scale, the probability that the next individual dying is a male \( aa \) is \( \lambda_1 k \Lambda^{-1} \), where

\[
\Lambda = \lambda_1 (k + r) + \lambda_2 (N - k - l - r - s) + \lambda_3 (l + s),
\]

[ 102 ]
with similar expressions for the other five kinds of individual. In order to obtain a non-trivial distribution we must make the differences between the \( \lambda \)'s tend to zero as \( N \) increases and we write \( \lambda_1 = 1 - g_1, \lambda_2 = 1, \lambda_3 = 1 + g_2 \), where \( g_1 = \gamma_1 N^{-1}, g_2 = \gamma_2 N^{-1} \) and \( \gamma_1, \gamma_2 \) are fixed quantities which may be positive or negative.

We can now confine ourselves to the embedded Markov chain, since the transition probabilities from any state depend only on that state, and the length of time which the system spends in any particular state without a death occurring will have a distribution which is asymptotically independent of the state, since the \( \lambda \)'s tend to equality.

We could have introduced selection in another manner by supposing that each zygotic individual has a varying capacity to produce gametes. This would almost certainly give the same results as the above assumption. If all the births occur simultaneously, as in Wright's model and the model of the following paper (part II), we would have to use this second model of selection, since we could not vary the lifecycle distribution of the zygotic individuals.

In previous papers completely random mating has been assumed. To relax this condition we introduce a certain degree of correlation between uniting gametes. This requires some care. Let \( p_M(a) \) and \( p_M(A) \) be the relative frequencies of \( a \) and \( A \) genes in the male gametic output, and \( p_F(a), p_F(A) \) similarly in the female gametic output. We suppose that the probabilities of matings between gametes are given by the expression in the cells of the \( 2 \times 2 \) table

<table>
<thead>
<tr>
<th>female a</th>
<th>male (a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_F(a) )</td>
<td>( (1-f) p_M(a) p_F(a) + \frac{1}{2} f (p_M(a) + p_F(a)) )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>female A</th>
<th>male (A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_F(A) )</td>
<td>( (1-f) p_M(A) p_F(A) + \frac{1}{2} f (p_M(A) + p_F(A)) )</td>
</tr>
</tbody>
</table>

Here \( f \) is a number satisfying \( 0 \leq f \leq 1 \). When \( f = 0 \) we clearly have random breeding and as \( f \) increases we have a greater and greater degree of positive assortative mating between gametes. Such non-randomness might arise from inbreeding or from assortative phenotypic mating between zygotes.

The difficulty of defining a table of this kind arises from the fact that we want to generalize the usual deterministic treatment in which \( p_M(a) = p_F(a) \) to a model which \( p_M(a) \) and \( p_F(a) \) are possibly quite different. If \( p_M(a) = p_F(a) = p = 1 - q \), the relative frequency of gametic matings of types \( aa, Aa \) and \( AA \) are \( (1-f) p^2 + fp, 2(1-f)pq \) and \( (1-f)q^2 + fq \). The above table appears to be the only simple way of generalizing this.

\( f \) may be given two interpretations. If \( p_M(a) = p_F(a) \) and the gametes \( a \) and \( A \) are scored 0 and 1, \( f \) is the product-moment correlation coefficient between uniting gametes. Alternatively we may suppose that a fraction \( 1 - f \) of all gametic matings are strictly at random, but that a fraction \( f \) provide \( aa \) and \( AA \) in proportions \( \frac{1}{2} (p_M(a) + p_F(a)) \) and \( \frac{1}{2} (p_M(A) + p_F(A)) \).

Although the above table appears to be the only simple one it suffers from some defects. The row totals are not equal to \( p_F(a) \) and \( p_F(A) \) and similarly for the column
totals. Moreover, if \( p_M(a) = 0, p_F(a) > 0 \), some \( a a \) will be found amongst the offspring which is biologically impossible. However, we shall prove that \( p_M(a) - p_F(a) \) converges in probability to zero so that these difficulties are not serious for \( N \) large.

A more curious fact is that \( f \) must be non-negative. If \( f \) were a negative number, it would be possible to find values of \( p_M(a), p_F(a) \) which would make the output of \( a a \) zygotes negative. To introduce a negative correlation between uniting gametes we should have to proceed differently. In this case we might suppose that the probabilities of \( a a, A a \) and \( A A \) zygotes in the next generation are

\[
(1-f) p_M(a) p_F(a),
\]

\[
(1-f) \{p_M(a) p_F(A) + p_M(A) p_F(a)\} + f,
\]

\[
(1-f) p_M(A) p_F(A),
\]

respectively, where \( 0 \leq f \leq 1 \). This is what we get by interchanging \( a \) and \( A \) in one sex but not in the other. Then \(-f\) could be regarded as a measure of negative association between uniting gametes. If \( f > 0 \), however, this leads to the unexpected result that as \( N \) becomes large \( p_M(a) \) and \( p_F(a) \) converge in probability to \( \frac{1}{2} \) and the probability distribution of gene frequency is degenerate. The reason for this becomes clear if we again compare the deterministic cases. Then with \( p = p_M(a) = p_F(a) \) and negative association we get \( (1-f) p^2, 2(1-f) pq + f \) and \( (1-f) q^2 \) for the probabilities of producing \( a a, A a \) and \( A A \). It is easy to verify that this changes the proportion of \( a \) gametes in the next generation towards the value \( \frac{1}{2} \), whereas in the corresponding deterministic model with positive association, the gene frequency is unaltered.

We can now consider the transition probabilities of the process. The state of the system is defined by four variables, \((k, l, r, s)\), and the matrix of transition probabilities would be impossibly complicated to write down. However, from any given state, at most twelve other states are accessible and we can consider these individually.

For convenience we introduce random variables defined by \( X_t = kN_1^{-1}, Y_t = lN_1^{-1}, W_t = rN_2^{-1}, Z_t = sN_2^{-1} \), where \((k, l, r, s)\) is the state of the system at time \( t \). Then the probabilities of death of the six types of individual are given by the following expressions:

- male \( a a \) \( N_1^{-1} \lambda_1 X_t \),
- male \( A a \) \( N_1^{-1} \lambda_3 (1 - X_t - Y_t) \),
- male \( A A \) \( N_1^{-1} \lambda_3 Y_t \),
- female \( a a \) \( N_2^{-1} \lambda_1 W_t \),
- female \( A a \) \( N_2^{-1} \lambda_3 (1 - W_t - Z_t) \),
- female \( A A \) \( N_2^{-1} \lambda_3 Z_t \).

Taking account of mutation the gene frequencies in the gametic output of the male and female parents are given by

\[
p_M(a) = \frac{1}{2} \{(1 - \alpha_1 + \alpha_2) + (X_t - Y_t) (1 - \alpha_1 - \alpha_2)\},
\]

\[
p_M(A) = \frac{1}{2} \{(1 + \alpha_1 - \alpha_2) - (X_t - Y_t) (1 - \alpha_1 - \alpha_2)\},
\]

\[
p_F(a) = \frac{1}{2} \{(1 - \alpha_1 + \alpha_2) + (W_t - Z_t) (1 - \alpha_1 - \alpha_2)\},
\]

\[
p_F(A) = \frac{1}{2} \{(1 + \alpha_1 - \alpha_2) - (W_t - Z_t) (1 - \alpha_1 - \alpha_2)\}.
\]
Write \( p(aa) \), \( p(Aa) \) and \( p(AA) \) for the probabilities that the newly formed individual will have genotypes \( aa \), \( Aa \) and \( AA \). Then using the above expressions for the probabilities of mating of the gametes, we find

\[
p(aa) = \frac{1}{2}(1 - f) \left\{ (1 - \alpha_1 + \alpha_2)^2 + (X_t - Y_t + W_t - Z_t) (1 - \alpha_1 - \alpha_2) (1 - \alpha_1 + \alpha_2) \right. \\
+ (X_t - Y_t) (W_t - Z_t) (1 - \alpha_1 - \alpha_2)^2 \left. + \frac{1}{2} f(2(1 - \alpha_1 + \alpha_2) + (X_t - Y_t + W_t - Z_t) (1 - \alpha_1 - \alpha_2) (1 - \alpha_1 + \alpha_2) \right. \\
- (X_t - Y_t) (W_t - Z_t) (1 - \alpha_1 - \alpha_2)),
\]

\[
p(Aa) = \frac{1}{2} (1 - f) \left\{ (1 - \alpha_1 + \alpha_2) (1 + \alpha_1 - \alpha_2) + (X_t - Y_t + W_t - Z_t) (1 - \alpha_1 - \alpha_2) (1 - \alpha_1 + \alpha_2) \\
- (X_t - Y_t) (W_t - Z_t) (1 - \alpha_1 - \alpha_2)^2 \right. \\
+ (X_t - Y_t) (W_t - Z_t) (1 - \alpha_1 - \alpha_2)^2 \left. + \frac{1}{2} f(2 (1 + \alpha_1 - \alpha_2) \right. \\
- (X_t - Y_t + W_t - Z_t) (1 - \alpha_1 - \alpha_2)),
\]

\[
p(AA) = \frac{1}{2} (1 - f) \left\{ (1 - \alpha_1 + \alpha_2)^2 - (X_t - Y_t + W_t - Z_t) (1 - \alpha_1 - \alpha_2) (1 + \alpha_1 - \alpha_2) \\
+ (X_t - Y_t) (W_t - Z_t) (1 - \alpha_1 - \alpha_2)^2 \right. \\
- (X_t - Y_t + W_t - Z_t) (1 - \alpha_1 - \alpha_2)).
\]

The system can move from the state \((k, l, r, s)\) into any one of the twelve states listed below with their corresponding transition probabilities. It can also remain in the state \((k, l, r, s)\), but it is not necessary to calculate explicitly the probability of this happening.

\[
\begin{align*}
(k - 1, l, r, s) & \quad \lambda_1 \Lambda^{-1} N_1 X_t p(Aa), \\
(k - 1, l + 1, r, s) & \quad \lambda_1 \Lambda^{-1} N_1 X_t p(AA), \\
(k + 1, l, r, s) & \quad \lambda_2 \Lambda^{-1} N_2 (1 - X_t - Y_t) p(aa), \\
(k, l + 1, r, s) & \quad \lambda_2 \Lambda^{-1} N_2 (1 - X_t - Y_t) p(AA), \\
(k + 1, l - 1, r, s) & \quad \lambda_3 \Lambda^{-1} N_1 Y_t p(aa), \\
(k, l - 1, r, s) & \quad \lambda_3 \Lambda^{-1} N_1 Y_t p(Aa), \\
(k, l, r - 1, s) & \quad \lambda_3 \Lambda^{-1} N_2 W_t p(AA), \\
(k, l, r - 1, s + 1) & \quad \lambda_3 \Lambda^{-1} N_2 W_t p(AA), \\
(k, l, r + 1, s) & \quad \lambda_2 \Lambda^{-1} N_2 (1 - W_t - Z_t) p(aa), \\
(k, l, r, s + 1) & \quad \lambda_2 \Lambda^{-1} N_2 (1 - W_t - Z_t) p(AA), \\
(k, l, r + 1, s - 1) & \quad \lambda_3 \Lambda^{-1} N_3 Z_t p(aa), \\
(k, l, r, s - 1) & \quad \lambda_3 \Lambda^{-1} N_3 Z_t p(Aa).
\end{align*}
\]

We now prove that \( X_t - W_t \) and \( Y_t - Z_t \) converge in probability to zero when \( N_1 \) and \( N_2 \) become large. In the rest of this paper we are going to consider the asymptotic behaviour of the system as \( N \) tends to infinity and we shall always assume that \( N_1 N_2^{-1} \) lies between fixed non-zero limits so that \( N, N_1 \) and \( N_2 \) are all of the same order and we can always replace an expression such as \( O(N_1^{-1}) \) by \( O(N^{-1}) \).

It is sufficient to prove that \( E(X_t - W_t)^2 \) tends to zero and we do this in exactly the same manner as in the previous paper (Moran 1958b). We have, when the system is stationary

\[
E(X_{t+1} - X_t + W_{t+1} - W_t + X_t - W_t)^2 - E(X_{t+1} - W_{t+1})^2 = 0
\]

\[
= 2E(X_{t+1} - X_t + W_{t+1} - W_t) (X_t - W_t) + E(X_t - W_t)^2.
\]
Enumerating the various cases given the state at time $t$ and multiplying through by $\Lambda$ we get (omitting the suffix $t$)

\[
N_1^{-1} E\{\lambda_1 X p(Aa) + \lambda_1 X p(AA) + \lambda_2 (1 - X - Y) p(aa) + \lambda_3 Y p(aa)\}
+ N_2^{-1} E\{\lambda_1 W p(Aa) + \lambda_1 W p(AA) + \lambda_2 (1 - W - Z) p(aa) + \lambda_3 Z p(aa)\}
+ 2E(X - W)\{- \lambda_1 X p(Aa) - \lambda_1 X p(AA) + \lambda_2 (1 - X - Y) p(aa) + \lambda_3 Y p(aa)\}
- 2E(X - W)\{- \lambda_1 W p(Aa) - \lambda_1 W p(AA) + \lambda_2 (1 - W - Z) p(aa) + \lambda_3 Z p(aa)\} = 0.
\]

This reduces to

\[
N_1^{-1} E\{\lambda_1 X p(Aa) + \lambda_1 X p(AA) + \lambda_2 (1 - X - Y) p(aa) + \lambda_3 Y p(aa)\}
+ N_2^{-1} E\{\lambda_1 W p(Aa) + \lambda_1 W p(AA) + \lambda_2 (1 - W - Z) p(aa) + \lambda_3 Z p(aa)\}
+ 2E(X - W)^2\{- \lambda_1 p(Aa) - \lambda_1 p(AA) - \lambda_2 p(aa)\}
- 2(\lambda_3 - \lambda_2) E(X - W) (Y - Z) p(aa).
\]

Since $\lambda_1 = 1 - \gamma_1 N^{-1}$, $\lambda_2 = 1$ and $\lambda_3 = 1 + \gamma_2 N^{-1}$ the first, second and fourth terms in this expression are all $O(N^{-1})$, whilst the third is asymptotically equal to $-2E(X - W)^2$ which must therefore also be $O(N^{-1})$.

In a previous paper (Moran 1958b) it has been shown that when $N$ increases the Hardy–Weinberg relation, which is a consequence of random mating, is satisfied with an error which tends to zero in probability. When $f$ is not zero we would expect that this relationship would be replaced by the convergence of

\[
(1 - f) \{4X_t Y_t - (1 - X_t - Y_t)^2\} - 2(1 - X_t - Y_t)
\]
to zero in probability, together with a similar result for $W_t$ and $Z_t$. This is just what we would expect if $X_t, p(aa)$ and $W_t, p( AA)$ converge to zero in probability which would be easy to prove if each generation was produced simultaneously as in Wright’s model. Here we have to proceed in a less direct manner along the lines of the previous paper.

Write $U_t = X_t - Y_t, V_t = X_t + Y_t$ and consider their joint moment generating function

\[
\Phi_t(\theta, \psi) = E \exp(\theta U_t + \psi V_t).
\]

Then

\[
\Phi_{t+1}(\theta, \psi) - \Phi_t(\theta, \psi) = 0
= E \exp(\theta U_t + \psi V_t) \{E[\exp[\theta(U_{t+1} - U_t) + \psi(V_{t+1} - V_t)] - 1]\}
= N_1^{-1} E \exp(\theta U_t + \psi V_t) \{\lambda_1 X_t p(AA) \exp(- \theta N_1^{-1} - \psi N_1^{-1})
+ \lambda_1 X_t p(AA) \exp(- 2\theta N_1^{-1}) + \lambda_2 (1 - X_t - Y_t) p(aa) \exp(\theta N_1^{-1} + \psi N_1^{-1})
+ \lambda_2 (1 - X_t - Y_t) p(AA) \exp(\psi N_1^{-1} - \theta N_1^{-1}) + \lambda_3 Y_t p(aa) \exp(2\theta N_1^{-1})
+ \lambda_3 Y_t p(AA) \exp(\theta N_1^{-1} - \psi N_1^{-1}) - \lambda_1 X_t p(Aa) - \lambda_1 X_t p(AA)
- \lambda_2 (1 - X_t - Y_t) p(aa) - \lambda_2 (1 - X_t - Y_t) p(AA) - \lambda_3 Y_t p(aa) - \lambda_3 Y_t p(AA)\].
\]
We expand the exponentials inside the main bracket and the absolute terms sum identically to zero. Substituting for the $\lambda$'s and retaining only terms of order $N^{-1}$ we obtain

\[
\Lambda^{-1} E \exp \left( \theta U_t + \psi V_t \right) \left\{ \theta \left[ -X_t p(Aa) - 2X_t p(AA) \right] \\
+ (1 - X_t - Y_t) \left[ p(aa) - p(AA) \right] + 2Y_t p(aa) + Y_t p(Aa) \right\} \\
+ \psi \left[ -X_t p(Aa) + (1 - X_t - Y_t) p(aa) + (1 - X_t - Y_t) p(AA) - Y_t p(Aa) \right].
\]

Since $\Phi_{t+1}(\theta, \psi) - \Phi_t(\theta, \psi) = 0$ identically the above expression must be not only $O(N^{-1})$, but also $o(N^{-1})$ uniformly for $\theta$ and $\psi$ in any bounded region. Substituting for $p(aa)$, $p(Aa)$ and $p(AA)$ and neglecting terms which are clearly $O(N^{-1})$ we find that

\[
\psi E \exp \left( \theta U_t + \psi V_t \right) \left\{ \frac{1}{2}(1 - f) \left[ 1 + (X_t - Y_t) (W_t - Z_t) \right] + f - (X_t + Y_t) \right\}
\]

is $o(N^{-1})$ uniformly for $\theta$, $\psi$ in any bounded region. Let $F_{N_1,N_2}(X_t, Y_t, W_t, Z_t)$ be the joint cumulative probability distribution. From the above we see that since $(W_t - Z_t)$ converges in probability to $(X_t - Y_t)$ then

\[
\int_0^1 \int_0^1 \int_0^1 \int_0^1 \exp \left( \theta U_t + \psi V_t \right) \left\{ \frac{1}{2}(1 - f) \left[ 1 + (X_t - Y_t)^2 + f - (X_t + Y_t) \right] \right\} dF_{N_1,N_2}(X_t, Y_t, W_t, Z_t)
\]

converges to zero uniformly for $\theta$, $\psi$ in any bounded region. This is the Fourier–Stieltjes transform of the function

\[
\int_0^X \int_0^Y \int_0^W \int_0^Z \left\{ \frac{1}{2}(1 - f) \left[ 1 + (x - y)^2 + f - (x + y) \right] \right\} dF_{N_1,N_2}(x, y, w, z),
\]

which must therefore converge to zero for all values of $X_t$, $Y_t$, $Z_t$ and $W_t$. The integral over any measurable set at which $\left| \frac{1}{2}(1 - f) \left( 1 - (x - y)^2 + f - (x + y) \right) \right| > \epsilon > 0$, say, must also tend to zero and thus

\[
\frac{1}{2}(1 - f) \left( 1 - (X_t - Y_t)^2 + f - (X_t + Y_t) \right)
\]

must converge in probability to zero.

Thus we have shown that the four variables $X_t$, $Y_t$, $W_t$ and $Z_t$ are such that the four expressions

\[
\begin{align*}
X_t - W_t \\
Y_t - Z_t \\
\frac{1}{2}(1 - f) \left( 1 - (X_t - Y_t)^2 + f - (X_t + Y_t) \right), \\
\frac{1}{2}(1 - f) \left( 1 - (W_t - Z_t)^2 + f - (W_t + Z_t) \right)
\end{align*}
\]

all converge to zero in probability.

Only three of these relationships are independent and the joint probability distribution of $(X_t, Y_t, W_t, Z_t)$ is concentrated asymptotically along a one-dimensional locus in a four-dimensional space.

To generalize Kolmogoroff's method using a diffusion equation to this situation would be very difficult, and it is easier to set up a differential equation satisfied asymptotically by the moment generating function of the gene frequency.
The relative frequency of \(a\)-genes in the whole population is \(\frac{1}{2} + \frac{1}{4}(k-l+r-s)N^{-1}\) for all \(N\). This turns out to be most inconvenient and we use instead
\[
\frac{1}{2} + \frac{1}{4}\{k-l\}N^{-1} + \{r-s\}N^{-1} = \frac{1}{2} + \{X_t-Y_t+W_t-Z_t\}
\]
which converges in probability to the true relative gene frequency. Moreover, it is simpler to begin by using the moment generating function of \(T = X_t-Y_t+W_t-Z_t\) which we write

\[
\phi_t(\theta) = E \exp \theta T_t.
\]

In the stationary state we have

\[
\phi_{t+1}(\theta) - \phi_t(\theta) = E\{\exp \theta (T_{t+1} - T_t) - 1\} \exp \theta T_t = 0.
\]

Enumerating the various cases we get (omitting the suffix \(t\))

\[
E \exp \theta T\{N_1\Lambda^{-1}[\lambda_1 X p(Aa) \exp (-\theta N^{-1}) + \lambda_1 X p(AA) \exp (-2\theta N^{-1})]
+ \lambda_3(1-X-Y)p(aa) \exp (\theta N^{-1}) + \lambda_3 Y p(Aa) \exp (\theta N^{-1})
- \lambda_1 X p(Aa) - \lambda_1 X p(AA) - \lambda_2(1-X-Y)p(aa)
- \lambda_3(1-X-Y)p(AA) - \lambda_3 Y p(Aa) - \lambda_3 Y p(Aa)\}
+ N_2\Lambda^{-1}[\lambda_1 W p(Aa) \exp (-\theta N^{-1}) + \lambda_3 W p(AA) \exp (-2\theta N^{-1})
+ \lambda_2(1-W-Z)p(aa) \exp (\theta N^{-1}) + \lambda_2(1-W-Z)p(aa) \exp (\theta N^{-1})
- \lambda_1 W p(Aa) - \lambda_1 W p(AA) - \lambda_2(1-W-Z)p(aa)
- \lambda_2(1-W-Z)p(AA) - \lambda_3 Z p(aa) - \lambda_3 Z p(Aa)\] = 0.

We expand the exponential terms inside the square brackets and consider the terms of order 1, \(N^{-1}\) and \(N^{-2}\), when \(N\) increases and \(N_1N_2^{-1}\) tends to a non-zero constant. The terms of order 1 vanish identically. We substitute for the \(\lambda_i\) and after some reduction, the coefficient of \(\theta\) in the expansion becomes

\[
E\Lambda^{-1} \exp \theta T\{(Y-X+W-Z) + 2(p(aa)-p(AA))g_2(Y+Z)+g_1(X+W)
- (g_1(X+W)-g_2(Y+Z))(p(aa)-p(AA))\}. 
\]

From previously given formulae we find that

\[
p(aa)-p(AA) = (\alpha_2-\alpha_1)+\frac{1}{2}(X-Y+W-Z)(1-\alpha_1-\alpha_2)
\]

and the above term becomes

\[
E\Lambda^{-1} \exp \theta T\{2(\alpha_2-\alpha_1) - (\alpha_1+\alpha_2)(X-Y+W-Z) + g_2(Y+Z)+g_1(X+W)
- (g_1(X+W)-g_2(Y+Z))(\alpha_2-\alpha_1)+\frac{1}{2}(X-Y+W-Z)(1-\alpha_1-\alpha_2)\}.
\]

Since \(\Lambda^{-1} = O(N^{-1})\) and \(\alpha_1, \alpha_2, g_1\) and \(g_2\) are all \(O(N^{-1})\) the above expression is \(O(N^{-2})\) and there are no terms of order \(N^{-1}\). We can easily verify, at this point, that if we had assumed negative association between uniting gametes, \(2(p(aa)-p(AA))\) would be equal to \((1-f)(X-Y+W-Z)+O(N^{-1})\), and the above expression would contain a term of order \(N^{-1}\) equal to

\[
E\Lambda^{-1} \exp \theta T\{f(X-Y+W-Z)\}.
\]
From this we could deduce that \((X - Y + W - Z)\) converges in probability to zero, and the gene frequenct to \(\frac{1}{2}\).

The terms of order \(N^{-2}\) will consist of two groups, one from the above expression, and one from the terms in the original expansion which contain \(\theta^2\). The latter are

\[
\frac{1}{2} \theta^2 E \Lambda^{-1} \exp \theta T [N_1^{-1} [\lambda_1 X p(AA) + 4 \lambda_1 X p(AA)]
+ \lambda_2 (1 - X - Y) p(aa) + \lambda_3 (1 - X - Y) p(AA) + 4 \lambda_3 Y p(aa) + \lambda_3 Y p(Aa)]
+ N_2^{-1} [\lambda_1 W p(Aa) + 4 \lambda_1 W p(AA) + \lambda_2 (1 - W - Z) p(aa)
+ \lambda_2 (1 - W - Z) p(AA) + 4 \lambda_3 Z p(aa) + \lambda_3 Z p(Aa)]
\]

Inserting the values of the \(\lambda\), substituting for \(p(aa), p(Aa)\) and \(p(AA)\), and neglecting terms whose contributions are \(O(N^{-3})\) we obtain

\[
\frac{1}{2} \theta^2 E \Lambda^{-1} \exp \theta T \left\{ \left( \frac{X + Y + W + Z}{N_1} + \frac{W + Z}{N_2} \right) - \left( \frac{X - Y + W - Z}{N_1} + \frac{W - Z}{N_2} \right) \right\} (X - Y + W - Z)
+ \left( \frac{1}{N_1} + \frac{1}{N_2} \right) (\frac{1}{2}(1 + f) + \frac{1}{2}(1 - f) (X - Y) (W - Z)) \right\}.
\]

Thus, the sum of all terms of order \(N^{-2}\) is equal to \(N^{-2}\) multiplied by

\[
\theta E N \Lambda^{-1} \exp \theta T \left\{ 2(\beta_2 - \beta_1) - (\beta_1 + \beta_2) (X - Y + W - Z)
+ B(Y + Z) + A(X + W) - \frac{1}{4}(A(X + W) - B(Y + Z)) (X - Y + W - Z)
+ \frac{1}{2} N \theta \left( \frac{X + Y + W + Z}{N_1} + \frac{W + Z}{N_2} \right)
- \frac{1}{2} N \theta \left( \frac{X - Y + W - Z}{N_1} + \frac{W - Z}{N_2} \right) (X - Y + W - Z) + \frac{1}{2} N \theta \left( \frac{1}{N_1} + \frac{1}{N_2} \right) (1 + f)
+ \frac{1}{2} N \theta \left( \frac{1}{N_1} + \frac{1}{N_2} \right) (1 - f) (X - Y) (W - Z) \right\}.
\]

This expression must converge to zero uniformly for \(\theta\) in any bounded region. We can now replace \(X - Y\) and \(W - Z\) by \(\frac{1}{2} T\) and \(X + Y\) by \(1 - \frac{1}{2}(1 - f)(1 - \frac{1}{2} T^2)\). This implies that we can also replace

\[
A(X + W) + B(Y + Z) - \frac{1}{4}(A(X + W) - B(Y + Z)) (X - Y + W - Z)
\]

by

\[
(A + B) (1 - \frac{1}{4} T^2) + \frac{1}{2}(1 - f) (1 - \frac{1}{4} T^2) (\frac{1}{2}(A - B) T - (A + B)).
\]

Moreover, \(NA^{-1}\) converges in probability to 1. Thus finally we have, writing

\[
M = N(N_1^{-1} + N_2^{-1})
\]

\[
E \exp \theta T [2(\beta_2 - \beta_1) + \frac{1}{2}(1 + f) (A + B) + \frac{1}{2} M \theta (1 + f)
+ [ - (\beta_1 + \beta_2) + \frac{1}{4}(1 - f) (A - B)] T
+ [ - \frac{1}{4}(A + B) + \frac{1}{2}(1 - f) (A + B) - \frac{1}{3} M \theta (1 + f)] T^2
+ [ - \frac{1}{16}(1 - f) (A - B)] T^3,
\]

This expression must converge to zero uniformly for \(\theta\) in any bounded region.
which converges uniformly to zero in any closed bounded region not containing the origin. In any such region it therefore follows that

\[
\epsilon_N = -\frac{1}{16}(1-f)(A-B)\Phi''(0) + \left\{ -\frac{1}{8}(1+f)(A+B) \\
-\frac{1}{8}M(1+f)\Phi'(0) + \left\{ -\beta_1 + \beta_2 + \frac{1}{4}(1-f)(A-B) \Phi'(0) \\
+ \left\{ 2(\beta_2 - \beta_1) + \frac{1}{2}(1+f)(A+B) + \frac{1}{2}M(1+f) \right\} \Phi(0),
\right.\] 

converges uniformly to zero. The proportion of a-genes in the population is

\[U = \frac{1}{2} + \frac{1}{4}T\] 

and writing \(\Phi(0) = E \exp^{\theta U} = \Phi(0) \exp^{\theta}\) we see that

\[-4(1-f)(A-B)\Phi''(0) + \left\{ 6(1-f)(A-B) - 2(1+f)(A+B) \\
-\frac{1}{8}M(1+f)\Phi'(0) + \left\{ -4(\beta_1 + \beta_2) + 2(1+f)(A+B) \\
-2(1-f)(A-B) + \frac{1}{2}M(1+f) \right\} \Phi'(0) + 4\beta_2\Phi(0)
\] 

tends to zero uniformly in any closed bounded region not containing the origin. If \(\Phi(0)\) therefore tends to a limit this limit must satisfy the third-order differential equation obtained by equating the above to zero. To solve the equation we guess a solution of the form

\[\phi(\theta) = \int_0^1 e^{\theta x + sx + tx^2} x^{d-1}(1-x)^{m-1} \, dx,
\]

which is suggested by Wright’s results. Substituting this in the differential equation, integrating by parts and equating to zero we find that one solution is given by

\[\phi(\theta) = \int_0^1 \exp \left\{ \theta x + 4x \left( \frac{(1+f)(A+B) - (1-f)(A-B)}{M(1+f)} \right) + 4x^2 \left( \frac{(1-f)(A-B)}{M(1+f)} \right) \right\}
\]

\[\times x^6 e^{\theta x(M(1+f))^{-1}(1-x)^{\beta(M(1+f))^{-1}} \, dx}, \quad (1)
\]

which satisfies the differential equation and is certainly the moment generating function of a distribution on the interval (0, 1) and satisfies the condition \(\phi(0) = 1\) for a suitable choice of \(K\).

We now have to show that \(\phi(\theta)\) does converge to this integral and thus that a limiting distribution does exist. The differential equation for \(\phi(N)\) can be written in the form

\[A_3 \phi''''(0) + (A_2 + B_2 \theta) \phi''''(0) + (A_1 + B_1 \theta) \phi'(0) + A_0 \phi(0) = \epsilon_N,
\]

where \(\epsilon_N\) tends to zero, when \(N_1, N_2\) tend to infinity in such a way that \(N(N_1^{-1} + N_2^{-1})\) tends to a constant \(M\). \(\phi(0)\) is the moment generating function of a distribution on (0, 1) and thus \(\phi(0) = 1\). Unlike the case of no selection considered in a previous paper, we do not know \(\phi'(0)\) and \(\phi''(0)\), or even whether they tend to a limit. A third-order differential equation requiring three constants to determine its solution, we thus cannot complete the argument by the methods used before. However, \(\phi'(0)\) and \(\phi''(0)\) are the first and second moments of \(U\) and are therefore bounded. It follows that out of any sequence of values of \(N_1, N_2\) such that \(N\) tends to infinity and \(N(N_1^{-1} + N_2^{-1})\) to \(M\), we can extract a subsequence for which \(\phi'(0)\) and \(\phi''(0)\) tend to limits. For this subsequence \(\phi(\theta)\) certainly converges to a function which is a
moment generating function of a distribution on \((0,1)\) and which satisfies the differential equation

\[
A_3 \phi''_1(\theta) + (A_2 + B_2 \theta) \phi'_1(\theta) + (A_1 + B_1 \theta) \phi'_1(\theta) + A_0 \phi_1(\theta) = 0
\]

everywhere (because \(\phi_1\) and its derivatives are continuous). We can therefore write

\[
\phi_1(\theta) = \int_{-\infty}^{1} e^{\theta x} dF(x),
\]

where \(F(x)\) is a non-decreasing function, continuous to the right, which has all its points of increase in the closed interval \((0,1)\) and the integral is defined to include the contribution of the jump at the point \(x=1\), if any such jump exists. Then since \(B_1 = -B_2\)

\[
\int_{-\infty}^{1} e^{\theta x} \left\{ A_3 x^3 + A_2 x^2 + A_1 x + A_0 - \theta B_2 x(1-x) \right\} dF(x) = 0
\]

identically in \(\theta\). Define

\[
G(x) = \int_{-\infty}^{x} \left\{ A_3 x^3 + A_2 x^2 + A_1 x + A_0 \right\} dF(x),
\]

where the integral includes the jump at \(x\), if any. \(G(x)\) is continuous to the right and constant for \(x \geq 1, x < 0\). Putting \(\theta = 0\) in (3) we see that \(G(1) = 0\). Then

\[
B_2 \theta \int_{-\infty}^{1} e^{\theta x} x(1-x) dF(x) = \int_{-\infty}^{1} e^{\theta x} dG(x)
\]

\[
= e^\theta G(1) - G(-\infty) - \theta \int_{-\infty}^{1} e^{\theta x} G(x) dx
\]

\[
= -\theta \int_{-\infty}^{1} e^{\theta x} G(x) dx.
\]

By the uniqueness theorem for Fourier–Stieltjes transforms it follows that

\[
B_2 \int_{-\infty}^{t} x(1-x) dF(x) = -\int_{-\infty}^{t} G(x) dx
\]

for all \(t\). But \(G(x)\), by (4), is bounded and the integral on the right is differentiable for all \(t\) and it follows that \(F(x)\) is differentiable in the open interval \((0,1)\). It then follows from (4) that \(G(x)\) is differentiable in the open interval \((0,1)\) so that (5) implies that \(F(x)\) can be differentiated twice in the open interval \((0,1)\). Next consider the limit of (5) as \(t\) tends to zero from above and suppose that \(F(x)\) has a non-zero jump of size \(d_0\) at \(x = 0\). Then from (4) the right-hand side of (5) will be of order \(A_0 d_0 t\), whilst the left-hand side will be not greater than \(B_2 t |F(t) - F(0)|\) because the jump in \(F(x)\) at \(x = 0\) does not contribute to the integral. But \(F(t) - F(0)\) tends to zero because \(F(x)\) is continuous to the right, and \(A_0 = 4\beta_2 \mp 0\). It follows that \(d_0\) is zero and by the symmetry of the process there can be no jump at \(x = 1\) so long as \(\beta_1 \neq 0\). Thus we can write \(\phi_1(\theta)\) in the form

\[
\phi_1(\theta) = \int_{-\infty}^{1} e^{\theta x} f(x) dx,
\]
and inserting this in (2) and integrating by parts we find that $\phi_1(\theta)$ must satisfy a first-order differential equation and there is only one arbitrary constant in the solution. This constant must be chosen to make $\phi_1(0) = 1$, and if we write $M(\theta)$ for the integral in (1)

$$K = M(0)^{-1}.$$  

Equation (1) is then the unique moment generating function of the limiting distribution (which therefore exists) of the gene frequency since it is completely independent of the particular subsequence of the $(N_1, N_2)$ used.

The limiting distribution therefore has a probability density proportional to

$$\exp 4x\left(\frac{(1+f)(A+B)-(1-f)(A-B)}{M(1+f)}\right)+4x^2\left(\frac{(1-f)(A-B)}{M(1+f)}\right)\times x^{8\beta_2[M(1+f)]^{1-1}(1-x)^{8\beta_2[M(1+f)]^{-1}}.}$$

This is of the same general form as that obtained by Wright for a population with simultaneous generations and random mating but the values of coefficients are different. If we put $A = B = f = 0$ so that mating is random and there is no selection, we obtain the same distribution as in a previous paper Moran (1958b). If we let $A = B = 0$ but $A \neq 0$, the selective value of the heterozygote is exactly intermediate between those of the homozygotes and the index of the exponential is linear instead of quadratic. Finally, the effect of non-randomness in the mating is to reduce the effective values of $\beta_1$ and $\beta_2$ since $f$ is greater than zero, and to replace $A-B$ by $(1-f)(1+f)^{-1}(A-B)$ whilst leaving $(A+B)$ unaffected. Thus, if the heterozygote is exactly intermediate between the homozygote, non-randomness in mating does not change the effect of selection, but does change the effective influence of the mutation rates. The former of these statements is in agreement with the results of Fisher (1922).

**References**