Approximate number of self-sterility alleles given the sum of their frequencies

(13) The approximate number of alleles given the sum of squares of their frequencies.

$$\alpha = S(p^2), \quad (21)$$

and allowing for the size of the population, N.

The number of alleles is given by

$$n = \sum_{1}^{N} \frac{2N}{a} \cdot \frac{1}{(N-1)!} \frac{(N-a)^{N-1}}{(1-2\alpha)^{N}} e^{-\frac{N-a}{1-2\alpha}} \quad (22)$$

in terms of the equilibrium distribution, when mutation balances extinction.

As the diagrams show (Genetical Theory of Natural Selection, Dover Publications Edition* p. 110) the frequency distribution among alleles of the frequency of representation in the population changes rapidly, and when $N\alpha^2$ exceeds 5, or even 2, approaches a compact unimodal curve.

Estimates of α can usually be obtained, but no direct estimates of n, consequently it may be of service to use the equilibrium condition to supply a provisional and approximate estimate of n.

Suppose

$$\left.\begin{aligned} x &= (N-a)/(1-2\alpha), \\ \text{then} \quad a &= N - x(1-2\alpha) \end{aligned}\right\} \quad 23$$

and the factor 2N/a may be expanded in the form

$$\frac{2N}{a} = 2\left\{1 + \frac{x}{N}(1-2\alpha) + \frac{x^2}{N^2}(1-2\alpha)^2 + \ldots\right\}. \quad (24)$$

Replacing the summation by integration with respect to x, as discussed on p. 107† we find

$$n = 2\left\{1 + (1-2\alpha) + \frac{N+1}{N}(1-2\alpha)^2 + \frac{(N+1)(N+2)}{N^2}(1-2\alpha)^3 + \text{ad inf}\right\} \quad (25)$$

The leading term is now

$$2\left\{1 + (1-2\alpha) + (1-2\alpha)^2 + \ldots\right\} = \frac{2}{2\alpha}, \text{ or } \frac{1}{\alpha}, \quad (26)$$

the term in $\frac{1}{N}$ is

$$\frac{2(1-2\alpha)^2}{N}\left\{1 + 3(1-2\alpha) + 6(1-2\alpha)^2 + \ldots\right\} = \frac{(1-2\alpha)^2}{4N\alpha^3}, \quad (27)$$

* Variorum edition, p. 295. —J.H.B.
† " " , p. 293.

while the next adjustment gives

$$\frac{2(1-2\alpha)^3}{N^2}\left\{2 + 11(1-2\alpha) + 35(1-2\alpha)^2 + 85(1-2\alpha)^3 + \dots\right\}$$

or

$$\frac{2(1-2\alpha)^3}{N^2}\left(\begin{array}{l} 3\left\{1 + 5(1-2\alpha) + 15(1-2\alpha)^2 + 35(1-2\alpha)^3 + \dots\right\} \\ -\left\{1 + 4(1-2\alpha) + 10(1-2\alpha)^2 + 20(1-2\alpha)^3 + \dots\right\}\end{array}\right)$$

$$= \frac{2(1-2\alpha)^3}{N^2}\left\{\frac{3}{32\alpha^5} - \frac{1}{16\alpha^4}\right\}$$

$$= (1-2\alpha)^3(3-2\alpha) \;/\; 16N^2\alpha^5 \qquad (28)$$

The expression sought is therefore given by the asymptotic series (29)

$$\frac{1}{\alpha}\left\{1 + \frac{(1-2\alpha)^2}{4N\alpha^2} + \frac{(1-2\alpha)^3(3-2\alpha)}{(4N\alpha^2)^2} + \frac{(1-2\alpha)^4(15-20\alpha+4\alpha^2)}{(4N\alpha^2)^3} + \dots\right\}$$

So if

$\alpha = .03$ $\qquad\qquad N = 1000$

$1 - 2\alpha = .94$ $\qquad\qquad 4N\alpha^2 = 3.6$

The third term is the smallest and convergence is unsatisfactory.

At $N = 3000$ the terms are

33.333
2.727
0.698
0.271
37.029

Making a total estimate of 37, with reasonable convergence at $N\alpha^2 = 2.7$.