57. *Statistical Studies on the Populations of Farm Animals. I* *Estimation of the Effective Population Size*

By Ken NOZAWA

Department of Animal Breeding, Faculty of Agriculture, Nagoya University

(Comm. by K. MASUI, M.J.A., April 12, 1957)

The effective population size, or the population number, is one of the most important conceptions in population genetics. This is defined as the number of breeding adults in an idealized population in which a random gamete has an equal probability of having come from any parent. In a population of a finite size the heterozygous loci of genes come to fixation by chance, generation after generation, independently from selection and other systematic evolutionary pressures. The effectiveness of the chance fixation of genes is inversely proportional to the effective size, namely

$$K = \frac{1}{2N},$$

where K is the rate of loss of heterozygosis per generation and N is the effective population size (Wright, 1931).

Wright (1931, 1938) derived the following formulae to estimate the effective population size (N) separately in some specific conditions of breeding structure:

$$N = 4N_m,$$

if the number of breeding males (N_m) is much less than that of breeding females (N_f), and

$$K = \frac{4N_0 - 2}{\sigma_k^2 + 2},$$

if N_0 parents contribute varying numbers (k) of gametes to the next generation, the mean of k being 2, and σ_k the standard deviation of numbers of gametes contributed by the individual parents. In breeding of farm animals, however, it is desirable to find a simple formula which makes it possible to calculate the effective size under the conditions of (i) the inequality of the numbers of breeding males and females, and (ii) the variability of the numbers of progeny of each parent. In this report the author will derive such a general formula which will include Wright's two formulae as its separate specific cases.

At first let us take into consideration the autosomal genes. We start with a parental population of N_m males and N_f females, and consider the distribution of numbers of gametes contributed by the individual members of parental population: in the male parents, the

i -th male contributes k_{mi} gametes to the next generation, mean, and standard deviation of k_{mi} being $\bar{k}_m$ and σ_{km} , respectively; in the female parents, the j -th female contributes k_{fj} gametes to the next generation, mean, and standard deviation of k_{fj} being $\bar{k}_f$ and σ_{kf} , respectively. If the progeny population is made up of N'_m males and N'_f females,

$$\sum_{i=1}^{N_m} k_{mi} = \sum_{j=1}^{N_f} k_{fj} = N'_m + N'_f.$$

When we sample two gametes randomly from the set of gametes coming from the parental population to the progeny population, the probability that these two have contributed by the same parent is

$$\begin{aligned} & \frac{\sum_{i=1}^{N_m} k_{mi}(k_{mi}-1) + \sum_{j=1}^{N_f} k_{fj}(k_{fj}-1)}{2(N'_m + N'_f) \{2(N'_m + N'_f) - 1\}} \\ &= \frac{N_m \sigma_{km}^2 + N_f \sigma_{kf}^2 + 2(\bar{k}_m + \bar{k}_f - 1)(N'_m + N'_f) - N_m \bar{k}_m^2 - N_f \bar{k}_f^2}{2(N'_m + N'_f) \{2(N'_m + N'_f) - 1\}}. \end{aligned}$$

If the effective population size is N , it is clear from the definition of this measure that the above probability is nearly $1/N$. If the difference of numbers of breeding adults between the parental and the progeny population is not large,

$$(1) \quad N = \frac{2(N_m + N_f) \{2(N_m + N_f) - 1\}}{N_m \sigma_{km}^2 + N_f \sigma_{kf}^2 + 2(\bar{k}_m + \bar{k}_f - 1)(N_m + N_f) - N_m \bar{k}_m^2 - N_f \bar{k}_f^2}.$$

This is a general formula to estimate the effective population size, in which both (i) the inequality of the numbers of breeding males and females, and (ii) the variability of the numbers of progeny of each parent are considered.

If the number of breeding males is much less than that of breeding females, as seen in some of the populations of farm animals, namely

$$N_m + N_f \doteq N_f$$

$$\bar{k}_m = N_f / N_m$$

$$\bar{k}_f \doteq 1$$

$$\bar{k}_m \gg \bar{k}_f,$$

and if the coefficients of variability of the numbers of gametes contributed by the members of the parents of the both sexes are $V_{km} (= \sigma_{km} / \bar{k}_m)$ and $V_{kf} (= \sigma_{kf} / \bar{k}_f)$, respectively, the formula (1) can be transformed into

$$N = \frac{4 N_f}{\frac{N_f}{N_m} (V_{km}^2 - 1) + (V_{kf}^2 - 1) + \frac{2 N_f}{N_m}}.$$

When N_f is sufficiently large, it may be considered that

$$\frac{2 N_f}{N_m} \gg |V_{km}^2 - 1|.$$

So, the second term of the denominator being neglected,

$$(2) \quad N = \frac{4N_m}{V_{km}^2 + 1}.$$

In addition, let us consider the sex-linked genes. Then

$$\sum_{i=1}^{N_m} k_{mi} = N'_f$$

$$\sum_{j=1}^{N_f} k_{fj} = N'_m + N'_f,$$

where k_{mi} and k_{fj} have to be considered as the numbers of sex chromosomes contributed by the members of the male and female parents, respectively. We get the following formula as corresponding to the formula (1):

$$(1') \quad N = \frac{(N_m + 2N_f)\{(N_m + 2N_f) - 1\}}{N_m\sigma_{km}^2 + N_f\sigma_{kf}^2 + (2\bar{k}_m - 1)N_f + (2\bar{k}_f - 1)(N_m + N_f) - N_m\bar{k}_m^2 - N_f\bar{k}_f^2}.$$

If the number of breeding males is much less than that of breeding females, the formula (1') becomes

$$N = \frac{4N_m}{V_{km}^2 + 1},$$

which is not different from the formula (2).

The formula (2) may be useful to the estimation of the effective population size of the existing animal breeding population. The remarkable fact shown in this formula is that in the population of farm animals in which the number of breeding males is much less than that of breeding females, the effective population size depends on the number of breeding males and the variability of the numbers gametes contributed by the individual male parent and is nearly independent to the corresponding quantities on the female side.

Finally it will be shown that Wright's two formulae mentioned above are included in the formulae (1) and (2) as specific cases. If the numbers of breeding males and females are equal, and if there is no sexual difference in the distribution of the numbers of gametes contributed by the members of parents, namely, if

$$N_m = N_f = N_0/2$$

$$\bar{k}_m = \bar{k}_f = \bar{k} = 2$$

$$\sigma_{km} = \sigma_{kf} = \sigma_k,$$

the formula (1) becomes

$$N = \frac{4N_0 - 2}{\sigma_k^2 + 2},$$

which is Wright's formula (1938).

In the case of the sex-linked genes,

$$N_m = N_f = N_0/2$$

$$\bar{k}_m = 1$$

$$\bar{k}_f = 2.$$

Then the formula (1') becomes

$$N = \frac{9N_0 - 6}{2(\sigma_{km}^2 + \sigma_{kf}^2) + 4}.$$

Let us consider a special condition where all the individual male parents produce accurately the same number of progeny, namely, if

$$V_{km} = 0,$$

the formula (2) becomes

$$N = 4N_m.$$

This is consistent to the result of derivation by Wright (1931).

The author is indebted to Dr. K. Ono and Mr. T. Shirao, both of the Department of Mathematics, Faculty of Science, Nagoya University, for their valuable criticisms. He is also grateful to Prof. K. Masui and Assist. Prof. K. Kondo of his department for their encouragement to the work.

References

- Wright, S. (1931): Evolution in Mendelian populations, *Genetics*, **16**, 97-159.
 — (1938): Size of population and breeding structure in relation to evolution, *Science*, **87**, 430-431.