Genet. Res., Camb. (1987), **49**, *pp.* 69–70 *Printed in Great Britain*

69

The general relationship between average effect and average excess

ALAN R. TEMPLETON

Department of Biology, Washington University, St Louis, Missouri 63130, U.S.A.

Summary

The average effect and average excess both measure the phenotypic effects of gametes in a population. A matrix notation is introduced that provides a general analytical solution for the average effects at a single locus with k alleles that can be solved for any population regardless of its genotype frequencies. This same notation also provides an easy way of deriving and generalizing to k alleles the well-known relationships between average effects and average excesses that exist under random-mating and regular deviations from Hardy–Weinberg equilibrium due to inbreeding.

Falconer (1985) has recently examined the relationships between R. A. Fisher's (1930) quantitative genetic measures of average effect and average excess for the two-allele case, and for deviations from Hardy–Weinberg genotype frequencies that are describable by the inbreeding coefficient, f. The purpose of this note is to show how some of Falconer's results can be generalized to an arbitrary number of alleles and to arbitrary deviations from Hardy–Weinberg. Moreover, the generalized relationships to be derived here allow the formulation of a general analytical solution for the average effect.

Both the average effect and the average excess represent ways of assigning a phenotypic effect to a haploid (gametic) genotype despite the fact that the phenotype is expressed only in diploid individuals. Phenotypic assignments to haploid genotypes play a critical role in quantitative genetic theory because diploid individuals pass on haploid gametes to the next generation in sexually reproducing populations. Consequently, when dealing with the transmission of diploid phenotypes from one generation to the next, it is essential to have some haploid measure of phenotypic transmissibility, since it is haploid gametes that constitute the genetic basis of transfer of phenotypic information across generations.

Fisher (1930) devised two such haploid phenotypic measures: average excess and average effect. Falconer (1985) gives several alternative definitions of each measure, but to avoid repetition I will only give one for each, following the definitions of Ewens (1979). The average excess of gamete type A_i is the conditional average phenotype of all individuals bearing at least one A_i gamete minus the overall population mean. This description can be expressed mathematically for a single locus with k alleles as follows. Let G_{ij} be the average phenotype for diploid individuals with genotype ij, $\overline{G}$ be the average phenotype for the population, $g_{ij} = G_{ij} - \overline{G}$ be the average phenotypic deviation of genotype ij, f_{ij} the frequency of the ordered genotype ij ($f_{ij} = f_{ji}$, such that $2f_{ij}$ equals the total frequency of genotype ij, ignoring allelic order), and p_i the frequency of allele i. The conditional frequency of genotype ij, given that one gamete involved in the fertilization is i, is $f(ij \mid i) = f_{ij}/p_i$. Then, the conditional average phenotypic deviation, given that allele i was involved in a fertilization event (the average excess of allele i, a_i) is

$$a_i = \sum_{j=1}^{k} f(ij \mid i) g_{ij}. \tag{1}$$

Note that equation (1) is a perfectly general expression for the average excesses of the alleles at a single locus. No assumption is made about the genotype frequencies being in Hardy–Weinberg equilibrium or deviating from Hardy–Weinberg equilibrium in a regular fashion.

The average effect is not a conditional mean deviation like the average excess, but instead it is the least-squares regression coefficient of the phenotypic deviations on the number of alleles of each type that make up the genotype. These least-squares regression coefficients, say α_i for allele i, are found by minimizing the equation

$$Q = \sum_i \sum_j f_{ij} (g_{ij} - \alpha_i - \alpha_j)^2. \tag{2}$$

Taking the partial derivative of Q with respect to each α_i and setting the resulting equations equal to zero yields the following equations:

$$a_i = \alpha_i + \Sigma_j f(ij \mid i)\,\alpha_j \quad i = 1,\dots,k. \tag{3}$$

It is now convenient to introduce the following matrix notation. Let $\tilde{a} = (a_1 \dots a_k)^T$ be the vector of average excess values, $\tilde{\alpha} = (\alpha_1 \dots \alpha_k)^T$ the vector of average effect values, I the $k \times k$ identity matrix with 1s on the diagonals and 0s elsewhere, and F the $k \times k$ matrix whose ijth element is $f(ij \mid i)$. Then the equations given in (3) can be expressed as

$$\tilde{a} = (I + F)\,\tilde{\alpha}. \tag{4}$$

A general analytical solution to the average effects follows immediately from equation (4); namely,

$$\tilde{\alpha} = (I + F)^{-1}\tilde{a}. \tag{5}$$

Note that equation (5) allows one to solve for the average effects for any set of genotype frequencies. Hence equations (4) and (5) give the basic general relationships between average effects and average excesses for a k-allele locus in a population with arbitrary genotype frequencies.

Equation (4) can be used to derive the well-known equality between average effects and average excesses under Hardy–Weinberg equilibrium. Let $\tilde{p} = (p_1 \dots p_k)$ be the row vector of allele frequencies. Then it follows from the definitions of average excess and average effects that $\tilde{p}\tilde{a} = \Sigma_j p_j a_j = 0$ and $\tilde{p}\tilde{\alpha} = \Sigma_j p_j \alpha_j = 0$ (Ewens, 1979). Under Hardy–Weinberg genotype frequencies, $F = P$ where P is the matrix each of whose rows is the vector $\tilde{p}$. From equation (4),

$$\tilde{a} - \tilde{\alpha} = F\tilde{\alpha} = P\tilde{\alpha} = \tilde{0} \tag{6}$$

where $\tilde{0}$ is the column vector whose elements are all 0s. This follows because each element of $P\tilde{\alpha}$ is given by $\tilde{p}\tilde{\alpha} = 0$. Hence average excesses and effects are identical under Hardy–Weinberg equilibrium.

Similarly, equation (4) can be used to generalize to k alleles the well-known relationship between average effect and average excess under regular deviations from Hardy–Weinberg equilibrium given by

$$f_{ii} = p_i^2 + p_i(1 - p_i)f = p_i^2(1 - f) + p_i f$$
$$f_{ij} = p_i p_j (1 - f), \quad i \neq j. \tag{7}$$

With these genotypic frequencies, $F = fI + (1 - f)P$. Hence,

$$\tilde{a} - \tilde{\alpha} = F\tilde{\alpha} = f\tilde{\alpha} + (1 - f)P\tilde{\alpha} = f\tilde{\alpha}. \tag{8}$$

It follows from (8) that $\tilde{\alpha} = \tilde{a}/(1 + f)$, which generalizes to k alleles the result given in equation (12) of Falconer (1985).

As can be seen from the above, the matrix formulations given in equations (4) and (5) are quite useful. First, they provide a general analytical solution for the average effects that can be solved for any population regardless of its genotype frequencies. Second, they provide an easy way of deriving and generalizing to k alleles the well-known relationships between average effects and average excesses under random-mating and under regular deviations from Hardy–Weinberg that can be described by an inbreeding coefficient.

References

Ewens, W. J. (1979). *Mathematical Population Genetics.* Berlin: Springer-Verlag.

Falconer, D. S. (1985). A note on Fisher's 'average effect' and 'average excess'. *Genetical Research* **46**, 337–347.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection.* Oxford: Clarendon Press.