LETTERS TO THE EDITOR

A GENERAL PROGRAM FOR MAXIMUM LIKELIHOOD ESTIMATION OF GENE FREQUENCIES

Dear Sir:

A general program for maximum likelihood estimation of gene frequencies (MAXIM) has been written in Fortran II. The program is completely general in the sense that it can handle any type of genetic system without the necessity of manually deriving and introducing particular equations into the program for specific runs. All the required equations are generated by the program from the input data which consist of the following:

1. A title card.
2. A card specifying the number of alleles and the number of phenotypes.
3. A card or cards specifying the observed number of each phenotype.
4. A card or cards specifying trial estimates (arbitrarily chosen) of each of the gene frequencies comprising the system under study.
5. A series of cards, one or more for each phenotype, which specifies by a particular code the genotypes comprising each specific phenotype.

The output consists of the values of the partial derivatives of the log likelihood equation, the maximum likelihood gene frequency estimates and their standard deviations, the expected numbers of each phenotype, a chi square for goodness of fit including the number of degrees of freedom, and the number of iterations performed to obtain the estimates to six decimal place accuracy.

In practice, the actual number of alleles and phenotypes that can be handled by the program is limited by the memory capacity of the computer. The present program utilizes 28,812 of the 32,500 core positions available in an IBM 1620 II computer with a memory capacity of 40,000 core positions. We have used ten for the maximum number of alleles, 30 for the maximum number of phenotypes, and 15 for the maximum number of genotypes included in any one phenotype. With these parameters, the program can be utilized to obtain gene frequency estimates in most commonly encountered situations. The program can be adapted to computers with larger capacity by changing some dimension statements.

The following example is based on Gm data for 187 Ainu of Japan (Steinberg, 1966). The following Gm alleles were postulated to explain the observed phenotypes: Gm^1, $Gm^{1,2}$, $Gm^{1,13}$, and Gm^2. Trial estimates of .25 were chosen for each allele, and after seven iterations the following results, based on eight significant figures, were obtained:

The investigations were supported in part by U. S. Public Health Service grant GM 07214 from the National Institutes of Health.

Allele	Frequency	σ	Phenotype	Observed	Expected
Gm^1	.539013	±.028722	Gm(1)	55	54.33
$Gm^{1,2}$	.038983	±.037528	Gm(1,2)	30	30.79
$Gm^{1,13}$	.317229	±.026616	Gm(1,2,13)	18	17.06
Gm^2	.104773	±.036082	Gm(1,13)	82	82.77
			Gm(2)	2	2.05
			TOTAL	187	187.00
			$\chi^2_{[1]} = .0894$;	$.75 < P < .90$	

The program has also worked satisfactorily with other sets of data, including one with seven alleles and 13 phenotypes, one of which had ten genotypes.

Further information and the program may be obtained by writing to the authors.

THADDEUS W. KURCZYNSKI*
ARTHUR G. STEINBERG
Department of Biology,
Western Reserve University,
Cleveland, Ohio 44106.

Received October 31, 1966.

REFERENCE

STEINBERG, A. G. 1966. Gm and Inv studies of a Hokkaido population: Evidence for a Gm^2 allele in the Ainu. *Amer. J. Hum. Genet.* 18: 459–466.

*Predoctoral trainee supported by U. S. Public Health Service grant 5T1 GM226 from the National Institutes of Health.