COUNTING METHODS IN GENETICAL STATISTICS

C.A.B.	Cedric A.B.	Smith
--------	-------------	-------

The Galton Laboratory, University College, London, Great Britain

Most parameters in genetical statistics are proportions. The frequency of a gene G is the proportion of genes at its locus which are of type G; the recombination fraction is the proportion of gametes, derived from a double heterozygote, which carry a recombination. The ideal way of estimating a proportion is by direct counting. Unfortunately that is often impossible in practice, either because of uncertainty, as when recessivity makes GG indistinguishable from Gg, or because of loss of cases, as in the estimation of the segregation ratio of a rare recessive, when matings Gg, by Gg which produce no recessive gg offspring cannot be identified, and so are omitted from the count. A method of overcoming these difficulties is to take a provisional value for the parameter to be estimated, use it to count the expected numbers of genes, recombinants, or segregants, including the "lost" cases if any, and so obtain an improved value: this is now taken as the provisional value, and the process repeated until the provisional and improved values agree. This method is closely related to maximum likelihood, and can be readily adapted to give the standard error of the estimate, and to test its homogeneity in a series of samples. These methods have already been described in part by Ceppellini, Síniscalco and Smith in the "Annals of Human Genetics" 20, 97-115, 1955, and a further explanation has been published in 1957 (ibid. 21, 254).

Acta genet., Vol. 7, No. 1, 1957