MAJOR GENES AND DEPARTURES FROM NORMALITY

T.W. Hancock

Waite Agricultural Research Institute, Glen Osmond, S.A. 5064

As most genes affecting quantitative traits are probably intermediate in effect between Mendelian and polygenic loci, neither form of genetic theory provides a particularly suitable means for analyzing data reflecting the effects of such genes. Depending on the frequency in the population and magnitude of the associated deviation, the presence of a major gene can lead to various forms of non-normality in the distribution, which may be skewed, leptokurtic, platykurtic or bimodal. While this is well known (Fisher *et al*, 1932; Mérat, 1968; Hammond and James, 1970; Piper, 1971; MacLean *et al*, 1976), there still remains a need for an effective procedure to discriminate between major loci and polygenic variation (Morton, 1974; Motulsky, 1978, Levine *et al*, 1978). The problem here is closely related to the more general statistical methods associated with the analysis of mixed distributions (i.e. Method of Moments (Hawkins, 1972) and Method of Maximum Likelihood (James, 1978)). As these tests are capable of producing multiple solutions (Hawkins, 1972) or contain singularities in the likelihood surface (James, 1978), it would seem undesirable to apply them to all data sets without first establishing evidence of deviation from normality.

Thus, a simple preliminary test for deviation from normality should provide an effective method of screening data sets. Any significant cases would then be assessed further to establish whether a major gene could be responsible for the non-normality and if so estimate its frequency and the mean and variance of individuals carrying it.

This approach differs from one reported by MacLean *et al* (1976) who have suggested the preliminary transformation of the data to remove skewness before analysis by maximum likelihood. As skewness is to be expected when the frequency of the major gene is low, transformation to remove it would be expected to reduce the power of the subsequent maximum likelihood analysis.

Computer simulation was used to investigate eight alternative test statistics to determine which is best suited to identifying the type of non-normality expected when a major gene is present in the population and secondly, to establish the minimum deviation which a major gene would have to cause before its presence could be consistently detected.

This demonstrated that the Modified Shapiro-Wilk statistic W' (Shapiro and Francia, 1972) provides the most versatile test. However, if prior knowledge suggested that the frequency of the major gene was near 0.5, then the 'lower-tailed' kurtosis statistic could be preferable. Unfortunately even these better tests require the major gene to deviate substantially from the general mean (i.e. for sample sizes of 500, deviation of two standard deviations is required before one can be reasonably sure of establishing non-normality).

These methods have been applied to the S.A. Merino sheep fleece weight selection experiment at Roseworthy Agricultural College (Mayo *et al*, 1969 and Hancock *et al*, 1979). No major genes were detected.

REFERENCES

FISHER, R.A., IMMER, F.R. and TEDIN, O. (1932) The genetical interpretation of statistics of the third degree in the study of quantitative inheritance. *Genetics* 17: 107-124.

HAMMOND, K. and JAMES, J.W. (1970) Genes of large effect and the shape of the distribution of a quantitative character. *Aust. J. Biol. Sci.* 23: 867-876.

HANCOCK, T.W., MAYO, O. and BRADY, R.E. (1979) Response to partial selection of clean fleece weight in South Australian strong wool Merino sheep. IV. Genetic parameters. *Aust. J. Agric. Res.* 30: 173-189.

HAWKINS, R.H. (1972) A note on multiple solutions to the mixed distribution problem. *Technometrics* 14: 973-976.

JAMES, I.R. (1978) Estimation of the mixing proportion in a mixture of two normal distributions from simple, rapid, measurements. *Biometrics* 34: 265-275.

LEVINE, J.D., GORDON, N.C. and FIELDS, H.L. (1978) Mixings of placebo analgesia. *The Lancet* 8091: 654-657.

MACLEAN, C.J., MORTON, N.E., ELSTON, R.C. and YEE, S. (1976) Skewness in commingled distributions. *Biometrics* 32: 695-699.

MAYO, O., POTTER, J.C., BRADY, R.E. and HOOPER, C.W. (1969) Response to partial selection of clean fleece weight in South Australian strong wool Merino sheep. I. Results of experiment. *Aust. J. Agric. Res.* 20: 151-167.

MÉRAT, P. (1968) Distribution de fréquencies interprétation du déterminisme génétique des caractères quantitatifs et réchérche de génes majeurs. *Biometrics* 24: 277-293.

MORTON, N.E. and MACLEAN, C.J. (1974) Analysis of family resemblance. iii Complex segregation of quantitative traits. *Amer. J. Hum. Genet.* 26: 489-503.

MOTULSKY, A.G. (1978) Medical and human genetics 1977: Trends and directions. *Amer. J. Hum. Genet.* 30: 123-131.

PIPER, L.R. (1971) Ph.D. Thesis, University of Edinburgh.

SHAPIRO, S.S. and FRANCIA, R.S. (1972) The approximate analysis of variance test for normality. *J. Amer. Stat. Assoc.* 67: 215-216.

* * * *