ROBUST ESTIMATION AND OUTLIER DETECTION FOR REPEATED MEASURES EXPERIMENTS.

R.M. HUGGINS

1. INTRODUCTION.

Standard methods of testing for repeated measures experiments and other mixed linear models are based on the likelihood function or least squares methods. These methods are known to be sensitive to model misspecification, and can be adversely affected by the presence of outliers.

There have been several attempts at constructing robust estimates of the parameters associated with such models, eg Rocke (1982), Fellner (1986), Huggins (1991a,b) but only the methods of Huggins (1991a,b) based on M-estimators allow ready computation of standard errors of the resulting estimates and the construction of hypothesis tests. Note that the parametric form of the covariance matrix in repeated measures experiments allows the adoption of a simpler approach than that of Maronna (1976), see also Carroll (1978), who takes a different approach to constructing M-estimates.

The approach used here also differs from that of Fellner (1986) who considered the robust estimation of variance components in mixed models. He constructed estimating equations, which are then solved iteratively, by trimming large residuals rather than constructing a robustified likelihood as is done here. The methods based on the t-distribution, as in Lange et al (1989), can be difficult to interpret in practice. Huggins (1991a) gives an example where the inclusion of an outlier in a dataset causes the degrees of freedom of the t-distribution to fall from 4945 to 13, which gives a misleading impression of the distribution of the bulk of the data. Methods based on the t-distribution have the further disadvantage that they downweight entire vectors of observations when it may only be one of the components of the vector which is abnormal.

The approach here supposes that one is interested in the behaviour of the "centre" of the data and that the estimates, particularly those for variance components, should be able to be interpreted in a multivariate normal framework, see Huggins (1991b) for a discussion of this. In particular one would like to obtain estimates very close to the maximum likelihood estimates if there are no outliers.