Multivariate Analysis and Its Applications
IMS Lecture Notes – Monograph Series (1994) Volume 24

ON CLUSTER REGRESSION AND FACTOR ANALYSIS MODELS WITH ELLIPTIC t ERRORS

BY BRAJENDRA C. SUTRADHAR
Memorial University of Newfoundland

This paper deals with a two-stage cluster sampling problem. At the first stage k clusters are drawn at random, and then at the second stage n p-dimensional correlated observations are chosen under each cluster, which may be linearly related to certain covariates. The data of this type can be represented by a cluster regression model with suitable distributional assumption for the errors. In this paper, it is assumed that the error vector of the linear model follow an np-dimensional elliptically contoured t-distribution. Many elliptical distribution theory results are developed in the literature under the assumption that these n p-dimensional errors are uncorrelated. But, in the case of cluster sampling, errors are usually assumed to be equicorrelated. We, therefore, introduce a suitable $np \times np$ covariance matrix for n p-dimensional errors which takes the common intra-cluster correlation into account. We then study the likelihood inferences for the regression parameters (coefficients of the covariates) of the (linear) cluster regression model. Maximum likelihood estimators (m.l.e.) of the regression parameters are found to be more efficient than the generalized least squares estimators for smaller values of the degrees of freedom parameter of the elliptically contoured t-distribution. The asymptotic $(k \to \infty)$ distribution of the m.l.e. of the regression coefficients is also given.

Further, a factor analysis model is studied. Based on the assumption that n p-dimensional observations are uncorrelated and they follow np-dimensional elliptically contoured t-distribution, we have developed Neyman's partial score test for testing the suitability of the number of factors. The test statistic has asymptotically χ^2 distribution with suitable degrees of freedom. Moreover, the test is asymptotically optimal.

1. Introduction. Elliptical distributions have been employed in two general approaches yielding somewhat different results. In one, p-dimensional

AMS 1980 Subject Classifications: 62H12, 62H15, 62H25, 62J05

Key words and phrases: Asymptotic chi-square distribution; clustered data; efficiency of the estimators; elliptically contoured t-errors; maximum likelihood estimation; Neyman's partial score test; testing covariance matrix.