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

UNBIASEDNESS OF A MULTIVARIATE OUTLIER TEST FOR ELLIPTICALLY CONTOURED DISTRIBUTIONS

BY JIAN-XIN PAN* AND XUE-REN WANG

Hong Kong Baptist College and Yunnan University

Under a class of elliptically contoured distributions, the likelihood ratio criterion (LRC) for detecting multiple outliers is established in this paper. It is shown that the LRC has the same form as in the normal case. Furthermore, the unbiasedness of the LRC is derived.

1. Introduction. Suppose that $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ is a random sample from some p-dimensional distribution $F_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma} > \mathbf{0}$ are the mean vector and the covariance matrix, respectively, and both are unknown. We are often required to detect whether there are outliers in the sample and which sample points are outliers. This can be reduced to a testing hypothesis problem as follows. If we choose the null model in which there are no outliers in the sample

$$H : \boldsymbol{x}_i \sim F_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \qquad i = 1, 2, \cdots n, \tag{1.1}$$

then a possible alternative model which may account for multiple outliers is the multivariate model with mean slippage

$$\begin{aligned} K : \quad & \boldsymbol{x}_i \sim F_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \qquad (i \notin I) \\ & \boldsymbol{x}_i \sim F_p(\boldsymbol{\mu} + \boldsymbol{a}_i, \boldsymbol{\Sigma}) \qquad (i \in I) \end{aligned} \tag{1.2}$$

where $I = \{i_1, i_2, \cdots, i_k\}$ is an index subset of $\{1, 2, \cdots, n\}$ with a fixed positive integer k and $\boldsymbol{a}_{i_1}, \cdots, \boldsymbol{a}_{i_k}$ are unknown mean slippage parameters.

For hypotheses (1.1) and (1.2), Siotaui (1959) and Wilks (1963) discussed the likelihood ratio criterion under the multivariate normal distribution $F_p = N_p$. In this paper, we extend their results to a class of elliptically contoured distributions. The null distribution of the likelihood ratio testing statistic for detecting the multiple outliers is shown to be a Wilk's distribution, which has the same distribution as in the normal case. Furthermore, the unbiasedness of the likelihood ratio test is derived. Under the assumption of an elliptically contoured distribution, Sinha (1984) provided a locally best invariant test for outliers based on multivariate sample kurtosis, which is useful for identifying

---

* Supported partially by the WAI TAK Investment and Loan Company LT.D. Research Scholarship of Hong Kong.

AMS 1991 Subject Classifications: Primary 62H15; Secondary 62H10.

Key words and phrases: Elliptically contoured distribution; likelihood ratio criterion; multiple outlier; unbiasedness of a test.