Nonparametric one-sided tests for multivariate data

Jong-Hwa Na
Chungbuk National University, Department of Statistics
(361-763) San 48 Gaeshin-dong, Heungduk-gu
Chong-ju, South Korea
cherin@cbucc.chungbuk.ac.kr

Hyo-Il Park
Chong-ju University, Department of Statistics
(360-764) 36 naeduk-dong, Sangdang-gu
Chong-ju, South Korea
hipark@chongju.ac.kr

Jeong-Suk Kim
Chungbuk National University, Department of Statistics
(361-763) San 48 Gaeshin-dong, Heungduk-gu
Chong-ju, South Korea
chastity@trut.chungbuk.ac.kr

1. Introduction

Let $X_1, X_2, ..., X_m$ and $Y_{m+1}, Y_{m+2}, ..., Y_N$ be two independent $d-$variate random samples from X and Y populations with sample sizes m and $n = N-m$ and with continuous distribution functions F and G, respectively. Since we are concerned with the location translation model, we assume that for all $x \in \mathbf{R}^d$, there is a $\theta \in \mathbf{R}^d$ such that

$$G(x) = F(x - \theta).$$

Based on this assumption, sometimes we are interested in testing the following hypotheses:

$$H_0 : \theta_1 \leq 0,\ \theta_2 \leq 0,\ \ldots,\ \theta_d \leq 0 \ v.s.\ H_1 : \text{at least one of } \theta_i'\text{s is strictly larger than } 0.$$

This is so-called one-sided testing problem for multivariate data. As an example, suppose that a laboratory has developed a medicine which may have effects on two symptoms simultaneously. One may draw a conclusion that this medicine is acceptable if it is effective for any one of two symptoms or for both. In this problem, the alternative under consideration can be restated as

$$H_1 : \text{at least one symptom may be cured.}$$

In spite of those versatile applicabilities of one-sided test procedures, the developments have not been so fruitful in nonparametric setting. Up to now, it seems that the main obstruction against the development of the one-sided nonparametric test procedures for mulitvariate data has been due to the non-availability of the tables for the $d-$variate normal distribution functions. For the bivariate case, Owen (1962) published a book which contains the bivariate normal distribution functions with 0 mean vector and unit variances by varying the values of correlation coefficient. However the tables are not sufficient even for the bivariate case since they can not contain all the values of the correlation coefficient. Nowadays we may overcome this inconvenience owing to the rapid progress of the ability of the statistical computations and so propose a new test procedures, which completely depends on computer to obtain the critical values for given significance levels.

In this paper, we propose one-sided nonparametric tests and apply the permutation principle to obtain the null distribution. For large sample case, we consider obtaining the tail probability of multivariate normal distribution with computer program and show an example. Also we consider asymptotic properties for our tests and modifications of our test statistics for various situations. Finally we compare the performances of our tests with the test based on the layer ranks through computer simulations.

REFERENCES

Bhattacharyya, G. K. and Johnson, R. A. (1970) : A layer rank test for ordered bivariate alternatives, *The Annals of Mathematical Statistics*, 41, 1296-1310.

Chatterjee, S. K. and Sen, P. K. (1964) : Non-parametric tests for the bivariate two-sample location problem, *Calcutta Statistical Association Bulletin*, 13, 18-58.

Park, H. I. and Desu, M. M. (1999) : A multivariate control median test. *Journal of the Statistical Planning and Inference*, 79(1), 123-139.

Puri, M. L. and Sen, P. K. (1971) : *Nonparametric methods in multivariate analysis*, Wiley, New York.

RESUME

We propose nonparametric one-sided test procedures for multivariate data. For small sample case, we may obtain the exact null distribution by applying the permutation principle. For large sample case, we consider the normal approximation for obtaining asymptotic tail probability or p-value. We illustrate our procedures through an example. Then we discuss the consistency and asymptotic power of our proposed test. Also we consider the cases when the directions of the unequal signs are reversed or are alternating in the hypotheses. Finally, we compare the performances of our test with those of the test based on the layer ranks through the empirical powers by computer simulations for some bivariate distributions.