

Selection problems based on ranked data

Takashi MATSUI

(Received January 9, 1997)

ABSTRACT. The purpose of this paper is to investigate some topics on the selection problems based on the vector and the combined ranks, with special reference to the structures of population parameters. We pay more attention to the vector rank statistics. First, by obtaining the joint distribution of the rank sums, exact results on the problems of selecting the best population based on the vector rank are given. Then, we give asymptotic results, for both the vector rank and the combined rank cases, using their respective moments. We put the main emphasis on the fact that these results are given with reference to the distributions (parameters) of the underlying populations. One of the open questions of the selection problems based on the ranked data lies in the determination of the *LFC*, though this problem has been discussed in several places. Thus, finally, under the assumption of the parametric configuration, some asymptotic results on *LFC* are obtained.

1. Introduction

In the analysis of experimental data, there are so many occasions to test the significance of k treatments. Analysis of variance technique is one of the statistical methods to cope with such situations. The problem is that even though we have an analytical result such as certain hypothesis being significant (or not significant), this may not be necessarily satisfactory to the experimenters. R. E. Bechhofer [3] states, in his first pioneering paper relating to ranking and selection, as follows: “Thus in an agricultural problem the hypothesis that several essentially *different* varieties of grain have the *same* (population) mean yield is an unrealistic one since it is obvious that if the varieties actually are different, the (population) mean yields also be different, and a sufficiently large sample will establish this fact at any preassigned level of significance. Moreover, should a significant result be obtained, the experimenter’s problem usually have just begun. For having established that the varieties are different he may now desire to select the one which is ‘best’. Here the best variety might be defined as the one having the *largest* (population)

1991 *Mathematics Subject Classifications.* 62F07, 62G30.

Key words and phrases. Combined rank, indifference zone, selection problem, subset selection, vector rank