RF Project 2241

FINAL REPORT

BY

THE OHIO STATE UNIVERSITY

RESEARCH FOUNDATION

1314 KINNEAR ROAD

COLUMBUS, OHIO 43212

To NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON, D.C. 20546

Grant No. NGR-36-008-040, Supplement 1

On NONPARAMETRIC RANKING AND SELECTION PROCEDURES

For the period 1 July 1967-30 September 1968

Submitted by Dr. M. Haseeb Rizvi
Department of Mathematics

NONPARAMETRIC RANKING AND SELECTION PROCEDURES

This report summarizes research performed under Grant No. NGR 36-008-040, Supplement 1 of NASA. The objective of this grant is research and development of nonparametric methodology in the relatively new area of statistical ranking and selection procedures.

Several problems have been studied and a number of papers written. The technical discussion of these follows.

A nonparametric selection procedure for the problem of selecting a subset of several populations which, with high probability, contains the one with the largest α-quantile has been considered and its optimal properties studied. The results are reported in a paper, written by the principal investigator (supervisor) and a co-author, which is entitled "Nonparametric procedures for selecting a subset containing the population with the largest α-quantile," Annals of Math. Statist., Vol. 38 (1967) pp. 1788-1803. Five reprints of this paper have already been transmitted to NASA.

Another nonparametric procedure based on order statistics has been proposed as a solution to the problem of selecting a subset of several populations which, with high probability, includes all populations that are better than a control population; here a population is considered to be better than the control if it has a smaller proportion of deficient items. The operating characteristics of this procedure are evaluated. Comparisons are made with two other competing nonparametric procedures in terms of asymptotic relative efficiencies; both Pitman and Bahadur type of efficiencies are considered. A related minimax procedure is also

studied. These results are given in the paper entitled "Nonparametric ranking procedures for comparison with a control" by the principal investigator and two co-authors; the paper is scheduled to appear in the December, 1968 issue of the Annals of Math. Statist., Vol. 39. Five preprints of this paper have been forwarded to NASA earlier and the reprints will be forwarded when they are received from the publishers.

Five preprints of the paper "On selection procedures based on ranks: Counter-examples concerning least favorable configurations" by M.H. Rizvi and G.G. Woodworth are being transmitted to NASA at this time. These results were presented by the first author at the annual meeting of the Institute of Mathematical Statistics held at Madison, Wisconsin during August 26-30, 1968. The paper has been submitted for publication. Specifically this paper is concerned with certain multiple-decision procedures based on rank sums which have been proposed for analyzing data in a one-way layout:

$$X_{ij} = \theta_i + \epsilon_{ij} , \quad i=1,\ldots,k, \; j=1,\ldots,n ,$$

where the errors $\{\epsilon_{ij}\}$ are independent, have the same known cumulative distribution function (cdf) F and where $\underset{\sim}{\theta} = (\theta_1,\ldots,\theta_k)$ is unknown. Two problems are considered:

I. Select the indices of the t largest θ-values.

II. Select a subset containing the index of the largest θ-value.

In problem I the experimenter sets a preassigned separation threshold $\delta^* > 0$ and preassigned probability threshold $P^* < 1$ and requires that the procedure he uses have the property that the probability of a correct selection is not smaller than P^* whenever the t largest θ-values are

at least δ^* larger than the rest of the θ-values. This problem might arise if there were k different batches of certain appliances available for use and one wanted to select the t best batches. In problem II the experimenter sets only the P^* value and requires that, with probability no smaller than P^*, the selected subset contains the index of the largest θ-value. This problem might arise in the first stage of screening different varieties of a certain instrument; one would want to reduce the number of varieties of the instrument which are to be submitted to further tests but at the same time be reasonably sure of not eliminating any variety which has great potentiality as to its effectiveness. Bartlett and Govindarjulu (1965), Lehmann (1963), Puri and Puri (1967,1968) and Woodworth (1965) have advanced certain procedures as solutions to problems I and II. In this paper we examine these procedures and show by means of specific examples that these procedures are in fact not solutions and should be used with great caution if they are used at all.

The work on numerical computations of the ranking integral involving noncentral chi-square (noncentral F) distribution and density functions is almost in the final phase. These computations will provide the tables needed for the problem considered by K. Alam and M.H. Rizvi in "Selection from multivariate normal populations," Annals Inst. Stat. Math., Vol. 18 (1966). pp. 307-318; this work was also partially supported by NASA under Grant No. NGR 36-008-040.

Some preliminary distribution-free results have been obtained for ranking components of the vactor parameter of M-ordered densities; these preliminary results were presented earlier by the principal investigator at the annual meeting of the Institute of Mathematical Statistics held at Washington, D.C. during December 27-30, 1967. Further results concerning the optimal decision-theoretic properties of the procedure for ranking the M-ordered densities are under investigation.

Some other nonparametric procedures closely related to the procedure proposed for the problem of comparison of several populations to a control population are also being investigated. The problem of constructing parametric as well as nonparametric confidence intervals for a ranked location (or scale) parameter is also under study. Work initiated for these problems will be more fully developed under a follow-on grant; a request for extension of the present grant in the form of Supplement 2 under the present grant number for a one-year period is being made to NASA at this time. It is hoped that the follow-on grant would lead to the development of a coherent theory of the nonparametric selection procedures and this could very well mean a substantial contribution to existing methodoloty of the statistical multiple-decision procedures.