

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

MULTIVARIATE RANKING AND SELECTION PROCEDURES  
WITH AN APPLICATION TO OVERSEAS TRADE

A thesis presented in partial fulfillment

of the requirements for the degree

of

MASTER OF ARTS

in

STATISTICS

at

MASSEY UNIVERSITY

Name: MARY SHARMILA MENDIS

Year: 1983

To my husband Rohan, our son Ruvan  
and  
the memory of our son Gerard

ABSTRACT

An overview of some recent work in the field of Ranking and Selection with emphasis on aspects important to experimenters confronted with Multivariate Ranking and Selection problems is presented. Ranking and Selection procedures fall into two basic categories. They are:

- 1) Indifference Zone Approach
- 2) Subset Selection Approach.

In these approaches, the multivariate parameters are converted to univariate parameters. Various procedures using these real valued functions are given for both the Indifference Zone Approach and the Subset Selection Approach.

A new formulation that has recently been developed which selects the best multivariate population without reducing populations to univariate quantities is also described. This method is a Multivariate Solution to the Multivariate Ranking and Selection problem.

Finally a real life problem pertaining to New Zealand's overseas trade is discussed in the context of Multivariate Ranking.

ACKNOWLEDGEMENTS

I wish to thank my supervisor Dr. Howard Edwards for his interest and guidance in the preparation of this thesis. Thanks also to Mrs. S. Carlyle for her excellent typing of the manuscript. Finally, I wish to thank my husband Rohan for his help, encouragement and support.

## TABLE OF CONTENTS

	<u>Page</u>
Abstract	iii
Acknowledgements	iv
List of Tables	ix
List of Computer Printouts	x
 <u>Chapter 1</u>	
1. Introduction	1
 <u>Chapter 2</u>	
2. Ranking and Selection	2
2.1 Possible Goals for Ranking and Selection Procedures	2
2.2 Approaches to Ranking and Selection Procedures	3
2.3 The Philosophy of the Indifference Zone Approach	3
2.3.1 Graphical Representation	5
2.4 The Philosophy of the Subset Selection Approach	6
2.5 Comparison of the Indifference Zone and the Subset Selection Approaches	7
 <u>Chapter 3</u>	
3. Multivariate Distributions	9
3.1 Definition of a Multivariate Distribution	9
3.2 Notation used in Multivariate Ranking and Selection	10
3.3 Multivariate Ranking and Selection	11
 <u>Chapter 4</u>	
4. The Indifference Zone Approach to Ranking and Selection of Several Multivariate Normal Populations	14
4.1 Selection in terms of the Mahalanobis Distance	14
4.1.1 Procedures studied by K. Alam and M.H. Rizvi (1966)	14
4.1.1.1 Case 1 $\Sigma_j$ known	16
4.1.1.2 Case 2 $\Sigma_j$ unknown	17

	<u>Page</u>	
4.1.2	Procedures studied by M.S. Srivastava and V.S. Taneja (1972)	18
4.1.2.1	Non Truncated Sequential Procedure	19
4.1.2.2	Truncated Sequential Procedure	21
4.2	Selection in terms of the Euclidean Distance	24
4.2.1	Procedures studied by M.S. Srivastava and V.S. Taneja (1972)	24
4.2.1.1	Case 1 $\Sigma$ known	25
4.2.1.2	Case 2 $\Sigma$ unknown	26
4.3	Selection in terms of Multiple Correlation Coefficients	27
4.3.1	Procedure investigated by M.H. Rizvi and H. Solomon (1973)	29
4.3.2	Procedure investigated by K. Alam, M.H. Rizvi and H. Solomon (1975)	31
4.4	Selection in terms of the Sum of the Bivariate Product - Moment Correlations	33
4.4.1	Procedure studied by Z. Govindarajulu and A.P. Gore (1971)	34
4.5	Selection in terms of the Coefficient of Alienation	35
4.5.1	Procedure proposed by R.M. Frischtak (1973)	36

## Chapter 5

5.	The Subset Selection Approach to Ranking and Selection of Several Multivariate Normal Populations	38
5.1	Selection in terms of the Mahalanobis Distance	38
5.1.1	Rules proposed by S.S. Gupta (1966)	39
5.1.1.1	Selection of a Subset containing the population with the largest $\theta_i$ , $\Sigma_1 = \Sigma_2 = \dots = \Sigma_k = \Sigma$ known	39
5.1.1.2	Selection of a Subset containing the population with the smallest $\theta_i$ , $\Sigma_1 = \Sigma_2 = \dots = \Sigma_k = \Sigma$ known	40
5.1.2	Procedures studied by S.S. Gupta and W.J. Studden (1970)	42
5.1.2.1	Selection of a Subset containing the population with the largest $\theta_i$ , $\Sigma_i$ not necessarily equal but known	42
5.1.2.2	Selection of a Subset containing the population with the smallest $\theta_i$ , $\Sigma_i$ not necessarily equal but known	43
5.1.2.3	Selection of a Subset containing the population with the largest $\theta_i$ , $\Sigma_i$ are different and unknown	43

	<u>Page</u>	
5.1.2.4	Selection of a Subset containing the population with the smallest $\theta_i$ , $\Sigma_i$ are different and unknown	45
5.1.3	Procedures considered by K. Alam and M.H. Rizvi (1966)	46
5.1.3.1	Selection of a Subset containing the population with the largest $\theta_i$ , $\Sigma_i$ not necessarily equal but known	46
5.1.4	Unsolved Procedures	47
5.1.4.1	Selection of a Subset containing the population with the largest $\theta_i$ , $\Sigma_1 = \Sigma_2 = \dots = \Sigma_k = \Sigma$ known	47
5.2	Selection in terms of the Generalized Variances	48
5.2.1	Rule proposed by M. Gnanadesikan and S.S. Gupta (1970)	48
5.2.1.1	Selection of a Subset containing the population with the smallest $ \Sigma_i $ based on the sample covariance matrices $S_i$ , $i = 1, \dots, k$	48
5.2.1.2	Approximations to the distribution of $Y_i$ when $p > 2$	49
5.2.2	Alternatives to Procedure R proposed by M.H. Regier (1976)	51
5.2.2.1	Procedure based on the Geometric Mean	51
5.2.2.2	Procedure based on the Arithmetic Mean	52
5.2.3	A comparison of the three procedures $R$ , $R_1$ and $R_2$	53
5.3	Selection in terms of the Multiple Correlation Coefficients	54
5.3.1	Procedures investigated by S.S. Gupta and S. Panchapakesan (1969)	55
5.3.1.1	Selection of $\theta_{[k]} = \rho_{[k]}$	55
5.3.1.2	Selection of $\theta_{[1]} = \rho_{[1]}$	56
5.4	Selection in terms of Measures of Association between two Subclasses of Variates	57
5.4.1	Procedure considered by S.S. Gupta and S. Panchapakesan (1969)	59
5.4.2	Procedure considered by R.M. Frischtak (1973)	59
 <u>Chapter 6</u>		
6.	The Multivariate Solution to the Multivariate Ranking and Selection Problem	61
6.1	The Multivariate Preference Selection Procedure $R_{mvp}$	62



	<u>Page</u>	
6.1.1	Selecting the best when $\Sigma_1 = \Sigma_2 = \dots = \Sigma_k = \sigma^2 I$ with $\sigma^2$ known	62
6.1.2	Selecting the best when $\Sigma_1 = \Sigma_2 = \dots = \Sigma_k = \Sigma$ , $\Sigma$ known	65
6.1.3	Selecting the best when $\Sigma_1, \dots, \Sigma_k$ known	67
6.1.4	Selecting the best when $\Sigma_1, \dots, \Sigma_k$ , unknown, unequal	69

### Chapter 7

7.	The Complete Ranking of Multivariate Populations	72
7.1	The Properties of the Multivariate Normal Distribution	72
7.1.1	Presentation of sample data consisting of n observations for each of the k populations each having p components	74
7.2	The Complete Ranking of k population means	75
7.2.1	Ranking of k normal populations according to their means when the common variance is known	75
7.2.2	Two stage procedure to rank k normal population means when the common variance is unknown	77
7.3	Ranking of Countries on their trade performance with New Zealand - An Application	78
7.4	Comments on the Application	85

### Chapter 8

8.	Conclusions	106
----	-------------	-----

	Bibliography	107
--	--------------	-----

LIST OF TABLES

	<u>Page</u>
T 4.1	Values of $t$ for $\lambda = 0$ 20
T 4.2	Values of $t$ 23
T 7.1	The probability of a correct complete ordering of $k$ normal populations with respect to means for given values of $\tau$ 87
T 7.2	Values of $h$ to determine the common sample size in the second stage for attaining $P^* = .95$ for a correct complete ordering of $k$ normal populations with respect to means when $n$ is the common size of the first stage sample 89

LIST OF COMPUTER PRINTOUTS

	<u>Page</u>
C 7.3.1 Results using Direct Data	90
C 7.3.2 Results when Data was transformed using Inverse Sine Conversion	94
C 7.3.3 Results when Data was transformed using Logit Conversion	98
C 7.3.4 Results when Data was transformed using Probit Conversion	102

## CHAPTER 1

### 1. INTRODUCTION

In the mid nineteen fifties attention began to be drawn to a new type of problem which does not fit into the framework of testing hypotheses and for which no proper statistical approach has been developed. In this type of problem it is not necessary to refute a null hypothesis which is clearly false but rather answer a different type of question which deals with selecting the best or with the ranking of alternatives. This field of study is called RANKING AND SELECTION THEORY.

A statistical Selection procedure uses a random sample from each population to select the best population. The same sample of data is used to order the populations in statistical Ranking procedures. In these procedures it can be asserted with a specified level of confidence that the Selection or Ranking made is correct.

Procedures for Selection and Ranking problems were pioneered by R. E. Bechhofer in 1954 using normality and equal known variance. In the ensuing years such procedures have been developed for more complex problems and in more realistic settings.

This thesis presents an overview of some recent work in this field with emphasis on aspects important to experimenters confronted with Multivariate Ranking and Selection problems. An example pertaining to overseas trade using Multivariate Ranking is also discussed.