Some Non-classical
Multivariate Distributions

FANG Hong-Bin

A thesis submitted in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

September 1998

Hong Kong Baptist University
Abstract

This thesis is devoted to the theory of multivariate distributions. The concentration is upon the two topics of essential statistical interests: (1) construction of new multivariate distributions for modelling data; (2) derivation of null and non-central distributions of statistics arising from robust statistical inference schemes. Some non-classical multivariate distributions are obtained based on the dependence pattern of random variables. The new non-central $t$- and $F$- distributions for high-dimensional mean tests with small sample size are derived.

It is well-known that the multivariate normal distributions have played a central role in the classical statistics. Because of the good properties of multivariate normal distributions, it is convenient to perform statistical inference using multivariate normal models for data sets. However, many data sets could not be well fitted by multivariate normal models. For some data sets, the dependence structures of the corresponding random variables are not as same as that of multivariate normal distributions even if the marginals are normal (cf. Cook and Johnson (1981)). On the other hand, some data sets have the dependence structures of normal distributions and non-normal marginals (cf. Krzysztofowicz and Kelly (1996)). So the need for studying multivariate non-normal distributions has been broadly recognised. Many multivariate non-normal distributions have been constructed and applied to some specific data in the past decades.

Yet it is difficult, if not impossible, to find a suitable multivariate distribution for modelling data. Many methods for specific data set have been proposed. Naturally, it is very important to construct various types of multivariate distributions that could be the most promising candidate for modelling
data. Emerged from different backgrounds, a lot of methods to construct multivariate distributions have been researched. As one of those methods, copula techniques has attracted more and more attention in the last decade. Copula, the dependence function of random variables, contains all the dependence information of the marginal random variables concerned. In terms of the study of dependence structures, it could be applied easily to construct multivariate distributions with desired properties.

In this thesis, the dependence patterns of random variables are emphasised. Since the copulas address all the dependence information of random variables filtering out the marginal distributions, the dependence patterns can be described precisely by copulas. The polynomial copulas, piecewise linear and quadratic copulas will be studied in Chapter 3. Especially, the bivariate distributions with “holes” in the domain will be studied. While the copula analysis is more or less self contained, even with the some copulas we can construct various multivariate distributions by embedding different marginal distributions.

As an extension of multivariate normal distributions, elliptically contoured distributions (ECD) have been extensively studied in the past two decades because they are a flexible class of distributions that are appropriate for modelling some data sets (Fang and Zhang (1990) and Fang, Kotz and Ng (1990)). Although many properties of ECD have been well explored, their dependence patterns have not been characterized. The copulas of ECD will be analysed in Chapter 4. It is shown that the subclasses of ECD have different dependence patterns. By means of the dependence structures, a much more general class of multivariate distributions, meta-elliptical distributions, is constructed.

The family of parameter distributions is very popular for statistical inference. With copula techniques, we will give a family of multivariate distributions of two parameters with free marginals. These distributions have very rich variations in their contours for different parameters. Hence, it is a quite resilient practical option for modelling data.

Chapter 6 is devoted to the derivation of non-central distributions. In multivariate statistics, usually it is difficult to handle the high-dimensional
data with small sample size. Many classical statistical tests are not feasible if the number of variables $p$ is greater than the number of observations $n$. Based on the theory of spherical matrix distributions (Fang and Zhang (1990)), Läuter et al. (1996) proposed some stable statistics for mean tests of normality population. These statistics can be used to manage high-dimensional data sets with small sample size. Naturally, the corresponding non-null distributions of these statistics are crucial. Based on the theory of zonal polynomials, the non-central distributions for the PC test will be derived.

Hong-Bin Fang
Department of Mathematics
Hong Kong Baptist University
224 Waterloo Road, Kowloon Tong
HONG KONG

September 25, 1998
Contents

Declaration ....................................................................................... i
Abstract ......................................................................................... ii
Acknowledgements ......................................................................... v
Notations and Abbreviations ............................................................ ix
List of Figures .................................................................................. xi

1. Introduction ................................................................................ 1
   1.1 Historical Review .................................................................. 1
       1.1.1 The Construction of Multivariate Distributions ................. 2
       1.1.2 Multivariate Distributions for Modelling Data ................. 10
       1.1.3 Statistical Inference of High-dimensional Data
            with Small Sample Sizes .................................................. 12
       1.1.4 Non-central Distributions .............................................. 15
   1.2 Motivation ............................................................................. 16
   1.3 Main Results of This Thesis .................................................. 18

2. Preliminary .................................................................................. 21
   2.1 Elliptically Contoured Distributions ...................................... 21
       2.1.1 The Properties of Elliptically Contoured Distributions ....... 21
       2.1.2 The Left-Spherical Distributions ................................... 24
   2.2 Zonal Polynomials .................................................................. 25
       2.2.1 Some Properties of Zonal Polynomials ............................. 25
       2.2.2 The Integral Formulae of Zonal Polynomials .................... 31

3. Copulas and Their Properties ...................................................... 34
   3.1 Copulas and Their Basic Properties ...................................... 35
   3.2 The Measures of Dependence .............................................. 44
   3.3 The Geometric Characterizations of Copulas .......................... 48