

ON SOME OPTIMAL TESTS IN FINITE MIXTURES : CONSTRUCTIONS AND APPLICATIONS

Chandranath Pal

Thesis submitted to Indian Statistical Institute, in partial
fulfillment of requirements for award of Doctor of Philosophy

November, 2000

Acknowledgements

It is my humble duty, at the very beginning, to express my deep gratitude to my supervisor, Professor Ashis SenGupta of the Indian Statistical Institute, Calcutta, for his constant encouragement and guidance throughout the work on this thesis. But for his active supervision with constructive comments and suggestions, it would have not been possible to complete the work. I also owe my indebtedness to Professor SenGupta for kindly giving his consent to include the materials of our joint papers in the thesis.

I thank the authority of the University of Kalyani for kindly sanctioning me study leave for a period of one year and four months during which I completed a part of this work.

I am also grateful to the Indian Statistical Institute authority for allowing me to work in the Institute on a full time basis during my study leave and thereby enabling me to fulfill the required residency attachment, after which I registered myself with the Institute for the Ph. D. degree.

Mr. S. Pal, Dr. S. Nandy of the I.S.I., Calcutta and Mr. C. Dawson of the UCSB, California, helped in the preparation of the diagrams and computations of some of the tables of this thesis. I am thankful to them. I also thank my colleagues of the Department of Statistics, University of Kalyani for various cooperations and help I received from them from time to time in connection with this work.

Finally I acknowledge my indebtedness to all my family members, particularly to my wife, for their enduring patience, encouragement and active support rendered towards the completion of this thesis.

Contents

Chapter/ Section		Page
1	Introduction and Summary	1
1.1	Introduction	1
1.2	Organization and summary of the thesis	5
2	Tests for Uniparameter Hypotheses in Regular Univariate Mixtures	12
2.1	Introduction	12
2.2	Nuisance parameters not present : LMP tests	13
2.3	Statistical curvature	15
2.4	Nuisance parameters present in the model	16
3	Tests for Uniparameter Hypotheses in Non-regular Univariate Mixtures	26
3.1	Introduction	26
3.2	Tests for the parameter	27
3.3	Tests for the mixing proportion	33
4	Tests in Mixtures of Distributions in Series Forms	36
4.1	Introduction	36
4.2	LMPI test in the SWS-CU mixture family	39
4.3	Monotonicity of the power function and consistency of the test	45
4.4	Simulation and computation	47
4.5	ρ known, a unknown : non-regular case	48
4.6	Examples	49

5	Tests for Uniparameter Hypotheses in Symmetric Multivariate Normal Mixtures		51
5.1	Introduction		51
5.2	LMP test in the SSMN mixture family		54
5.3	Statistical curvature		55
5.4	Exact power computations		58
5.5	Asymptotic distribution of the test statistic		59
5.6	SMN mixture family : the LMPS tests		60
5.7	Monotonicity of the power functions		62
5.8	Asymptotic normality and consistency of the tests		65
6	Higher Order $C(\alpha)$ Tests with Applications to Mixtures		68
6.1	Introduction		68
6.2	Non-existence of an optimal $C(\alpha)$ test		69
6.3	Second and higher order $C(\alpha)$ tests		70
6.4	Testing in multivariate models		74
7	The P^3 , L-optimal and Estimating Function Approaches		78
7.1	Introduction		78
7.2	The P^3 approach and L-optimal tests		80
7.3	Integrated likelihood approach and the optimal P^3 test		83
7.4	Mixtures in reliability analysis and L-optimal tests		84
7.5	Estimating function approach		88
7.6	Examples		91
7.7	L-optimal tests in the presence of nuisance parameters		93

8	Locally Bahadur Optimal Tests in Mixtures		98
8.1	Introduction		98
8.2	Local Bahadur optimality of a test		99
8.3	A locally optimal test in mixtures from the one-parameter exponential family		100
8.4	Examples		102
9	Tests for Multiparameter Hypotheses in Possibly Multicomponent Mixtures		106
9.1	Introduction		106
9.2	The LMMPU test		107
9.3	LMMPU tests in mixtures of univariate one-parameter distributions		108
9.4	LMMPU tests in multivariate mixtures		111
9.5	P^3 tests for multiparameter hypotheses in multivariate mixtures		113
9.6	Illustrative example		116
10	An Overview — Directions for Further Research		117
11	Computational Support and Results		121
11.1	Some selected computer programs		121
11.2	Tables		127
11.3	Diagrams		133
Appendices			135
Bibliography			142

Chapter 1

Introduction and Summary

1.1. Introduction

Mixtures of distributions are now-a-days playing very important roles in both theoretical and applied statistics. There is an abundance of real-life situations where mixture distributions are being extensively used for modelling data and drawing inference. Several books and monographs on mixtures have so far been published, *e.g.*, Everitt and Hand (1981), Titterington *et al.* (1985), McLachlan and Basford (1988), McLachlan (1997) etc., which cover a wide area on different aspects of mixture distributions and their applications. The book by Titterington *et al.* (1985, pp. 16-21), in particular, contains a comprehensive list of references on direct applications of finite mixtures in different fields of everyday life (see also Everitt, 1985; Titterington, 1997). However, while the literature on the estimation aspect of mixture distributions is quite rich, not much work seems to have been done on optimal testing problems.

One of the most important issues in finite mixtures in general and two-component mixtures in particular is that of testing for the hypothesis of ‘no mixture’ against a mixture alternative. The likelihood ratio test (LRT), even for a two-component mixture, is very cumbersome and regularity conditions do not hold for $-2 \log \lambda$ to have the usual asymptotic null distribution as that of a χ^2 with degrees of freedom equal to the difference in the number of parameters in the two hypotheses. Consider, for example, the mixture density

$$g(x|\mathbf{p}, \boldsymbol{\theta}) = \sum_{i=1}^r p_i f_i(x, \boldsymbol{\theta})$$

where $\mathbf{p} = (p_1, p_2, \dots, p_r)$; $p_i \geq 0 \forall i$ and $\sum_{i=1}^r p_i = 1$; $\boldsymbol{\theta}$ denotes a vector of unknown parameters $\theta_1, \theta_2, \dots, \theta_k$ and f_i s are the component densities. In the particular case of $r = 2$, testing the null hypothesis of ‘no mixture’

can be accomplished by testing whether $p_1 = 0$, which is on the boundary of the parameter space (Self and Liang, 1987; Feng and McCulloch, 1992), resulting in a breakdown in the standard regularity conditions.

When the mixture is composed of two known but general univariate densities f_1 and f_2 in unknown proportions p and $1 - p$ respectively, Titterton *et al.* (1985, pp. 152-153) considered the problem of testing $H_0 : p = 1$ against $H_1 : p < 1$ and showed that asymptotically, under H_0 , $-2 \log \lambda$ is distributed as $\{\max(0, Y)\}^2$ where Y is a standard normal variable.

To demonstrate the failure of the likelihood asymptotics in mixtures, Hartigan (1985) considered a mixture of two univariate normals $N(0, 1)$ and $N(\theta, 1)$ in unknown proportions $1 - p$ and p and showed that for testing $H_0 : \theta = 0$ against $H : \theta \neq 0$, the supremum of the loglikelihood, L_n , approaches ∞ in probability as $n \rightarrow \infty$.

Ghosh and Sen (1985) derived a general form of the asymptotic distribution of the LRT in 'strongly identifiable' mixtures. In the example of two-component univariate normal mixtures with unknown but identifiable means μ_1 and μ_2 , unknown mixing proportion p and common known variance, they showed that for testing $H_0 : p = 1$, the asymptotic distribution of $-2 \log \lambda$ is that of a certain functional

$$\left[\left\{ 0, \sup_{\mu_2} Y(\mu_2) \right\} \right]^2$$

provided μ_2 lies in a compact set, where $Y(\cdot)$ is a Gaussian process with zero mean and covariance kernel depending on the true value of μ_1 under H_0 and the variance of $Y(\mu_2)$ is unity for all μ_2 (see McLachlan and Basford, 1988, p. 22). Later Berdai and Garel (1996) extended this result to the case of unknown variance and obtained some tables by simulation.

Considering a mixture density of two k -variate normal distributions of the form

$$g(\mathbf{x}|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \Sigma) = pf(\mathbf{x}|\boldsymbol{\theta}_1, \Sigma) + (1 - p)f(\mathbf{x}|\boldsymbol{\theta}_2, \Sigma)$$

with known p , Goffinet *et al.* (1992) derived the asymptotic null distribution of the test statistic

$$T(\mathbf{X}) = 2 \left\{ \sup_{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \Sigma} \sum_{i=1}^n \log g(\mathbf{X}_i|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \Sigma) - \sup_{\boldsymbol{\theta}_1, \Sigma} \sum_{i=1}^n \log f(\mathbf{X}_i|\boldsymbol{\theta}_1, \Sigma) \right\}$$

for testing $H_0 : \theta_1 = \theta_2$ against $H_1 : \theta_1 \neq \theta_2$ corresponding to $k = 1$ (Theorem 1) and with known Σ corresponding to $k = 2$ (Theorem 2). For the last case they remarked ‘When Σ is unknown, the likelihood equations are very complex with a large number of terms, and no analytical result was found’.

The recent article by Titterton (1997) contains many more references on the LR approach of testing in mixtures. To avoid the above difficulties regarding huge computations and complicated asymptotic distributions, some modifications of the LR approach have been tried later. Furman and Lindsay (1994) worked with a mixture density

$$F_Q(x) = \sum_{i=1}^{\nu(Q)} p_i n(x|\mu_i, \sigma_0)$$

where $n(x|.,.)$ is the normal density and $\nu(Q)$ is the unknown number of components in the mixture. For testing $H_0 : \nu(Q) = r$ against $H_1 : \nu(Q) = r + 1$, they considered a test statistic, called the pseudo likelihood ratio statistic, of the form

$$L_r^* = 2 \log[\hat{g}_{r+1}/\hat{g}_r]$$

where $g_s = \sum_{i=1}^s p_i n(x|\mu_i, \sigma_0)$ and $\hat{g}_r, \hat{g}_{r+1}$ are obtained using the moment estimators of the parameters instead of the maximum likelihood estimators. Note that in the particular case of $r = 1$, H_0 is just the hypothesis of ‘no mixture’. Furman and Lindsay (1994) also carried out some simulations to study the power etc. of this test.

Another type of modification was considered by Chen (1994) who proposed a test for the null hypothesis $H_0 : g_1(x) = \sum_{i=1}^{k-1} p_i f(x, \theta_i)$ against the alternative $H_1 : g_2(x) = \sum_{i=1}^k p_i f(x, \theta_i)$ based on a statistic

$$\lambda_n = \sum_{j \in s_1} \log \hat{g}_1(x_j) - \sum_{j \in s_2} \log \hat{g}_2(x_j)$$

where s_1 and s_2 are two disjoint subsets of $\{1, 2, \dots, n\}$. Here also $k = 2$ reduces H_0 to the hypothesis of ‘no mixture’. The limiting null distribution

of λ_n , unlike the standard results, was seen to be normal. There is, however, a little bit of subjectivity regarding the choice of the partitions s_1 and s_2 .

Chernoff and Lander (1995) considered a mixture of two binomial distributions with p.m.f.

$$g(x|p) = p\text{Bi}(x|k, \pi) + (1 - p)\text{Bi}(x|k, \frac{1}{2})$$

with unknown p and π and derived the asymptotic null distribution of the LRT, under suitable reparametrizations, for the hypothesis that the population is a single binomial $\text{Bi}(k, \frac{1}{2})$.

All the above modified tests are still very much computation-intensive and moreover nothing is known about their optimality properties. The LR tests in general are known to be Bahadur optimal (Bahadur, 1960; 1967, Bahadur and Raghavachari, 1972) under certain conditions. In mixtures, however, verification of these conditions is extremely difficult and hence one can hardly ascertain such an optimality property.

Optimal tests, in the spirit of the Neyman-Pearson theory, can be constructed and in many situations they are quite simple and enjoy nice properties, *e.g.*, *monotonicity of the power functions, consistency etc.* While a lot of work has been done on the LR approach of testing in mixtures, works on optimal tests do not appear much in the literature. One of the main difficulties in constructing exact optimal tests is the presence of nuisance parameters in the mixture model, because then the problem, in many cases, cannot be reduced by any of the principles of sufficiency, invariance or similarity. Since uniformly best tests do not exist, in general, for mixtures, attempts have been made to construct locally best (LB) tests. Also in many cases it is more important but difficult to detect small departures from the null hypothesis since large departures can be detected quite easily by any reasonable test. Moreover in many distributional models, in general, and certain reliability models like bivariate Inverse Gaussian (Kocherlakota, 1986), in particular, locally most powerful (LMP) tests and/or their extensions in large samples (*e.g.*, $C(\alpha)$ tests) under the presence of nuisance parameters can be conveniently derived and are very much useful (see, *e.g.*, Kocherlakota and Kocherlakota, 1985).

The papers of Durairajan and Kale (1979, 1982) are some of the earlier works on the aspect of exact optimal testing in mixtures. Durairajan (1980), in his unpublished Ph. D. thesis, explored many of the optimal

testing problems in two-component mixtures and constructed LMP, LMP unbiased (LMPU), LMP invariant (LMPI) and LMP similar (LMPS) tests – the last two cases were considered when the model contained certain ‘structural parameters’, *e.g.*, location, scale etc.. Cabral (1991) considered asymptotically LMPU tests for a hypothesis of the form $H_0 : p = p_0$ against $H_1 : p \neq p_0$ where $p_0 \in (0, 1)$.

Failure of the existence of the above exact optimal tests in any case leads to construct optimal $C(\alpha)$ tests of Neyman (1959) or its different generalizations discussed in, *e.g.* Buehler and Puri (1966), Moran (1970) and Hall and Mathiason (1990).

There are, however, situations where none of the above exact or asymptotically optimal tests can be found directly because of certain types of non-regularities arising in the models. One type of non-regularity is the fact that the score function at the null value of the parameter vanishes identically in x , so that usual LMP tests cannot be found. Another type of non-regularity is attributed to the presence of nuisance parameters only under the alternative. In the later case, Neyman’s $C(\alpha)$ tests are not applicable and one has to follow the approach of Davies (1977, 1987).

The present thesis is an attempt to explore some optimal testing problems for the hypothesis of ‘no mixture’ in finite mixtures of probability distributions, although emphasis is laid on two-component mixtures. Both exact as well as asymptotically optimal tests where exact optimal tests do not exist or are very difficult to construct are derived and critically appraised. We consider both univariate and multivariate models and also uniparameter and multiparameter hypotheses and proceed, as far as practicable, in a logical way with different combinations of all these situations. A locally Bahadur optimal test is also proposed for mixtures of distributions from the one-parameter exponential family.

1.2. Organization and summary of the thesis

The thesis is organized through eleven Chapters and seven Appendices with contents as summarized below.

In Chapter 2 we restrict our attention to two-component ‘regular’ mixtures of univariate distributions and consider the problem of optimal testing

for uniparameter hypotheses. We briefly review the constructions of the LMP tests for the mixing proportion when the component densities are completely specified and also of the LMP (for one-sided alternatives) and the LMPU (for two-sided alternatives) tests for the parameter θ of the mixture distribution when the mixing proportion p is known. The aspect of statistical curvature as introduced by Efron (1975), used to impart an initial evaluation of the performance of an LMP test, is also considered. We next consider the case when, apart from θ , one or more nuisance parameters are present and prove a general result on the existence and construction of an LMP similar unbiased (LMPSU) test in mixtures. We finally take up three examples of mixtures of normal distributions useful in a wide variety of situations, *e.g.*, human genetics (Elston *et al.*, 1974), remote sensing (Odell, 1976), dental experiments (Marks and Rao, 1979) etc. and attempt to derive the unconditional LMPSU tests using the above theorem. Making simple reparametrizations the models can be reduced to ones involving location, scale and location-scale parameters which can be eliminated by similarity or by invariance. As we note, one of the models admits unique unconditional LMPSU test while for the other two, such a test does not exist. We proceed to derive LMPIU tests for these models. It is further noted that one of the examples does not even admit any unique LMPIU test. We defer this example to a subsequent Chapter where we derive the optimal third order $C(\alpha)$ test.

Chapter 3 deals with some univariate ‘non-regular’ mixture distributions, *i.e.*, distributions for which the score function vanishes identically in x so that the usual LMP tests cannot be constructed and further Efron’s statistical curvature cannot be computed in the usual way. Another type of non-regularity arises when nuisance parameters are present only under the alternative, so that Neyman’s approach is no longer applicable. We note that the first type of nonregularity can often be taken care of by suitable reparametrizations while the other can be tackled by the technique of Davies (1977, 1987).

The general model we consider here is (SenGupta and Pal, 1993a)

$$g(x|p, \mu, \sigma^2) = pN(x|\mu, \sigma^2) + (1 - p)N(x|-\mu, \sigma^2)$$

where $N(x|\mu, \sigma^2)$ is the p.d.f. of a normal distribution with mean μ and variance σ^2 . Note that $0 \leq p \leq 1, \sigma^2 > 0$ and μ is assumed to belong

to $[0, \infty)$ to ensure identifiability. We first consider the problem of testing $H_0 : \mu = 0$ against $H_1 : \mu > 0$ and construct locally most powerful (LMP) or $C(\alpha)$ tests under different situations concerning knowledge of the other parameters p and σ^2 . We note that the above non-regularities appear under different situations and these are tackled by appropriate techniques mentioned above. The role of Efron's statistical curvature (Efron, 1975) in this non-exponential family is also exploited. We derive the LMPS tests wherever possible.

The problem of testing for the mixing proportion p in the above model is then considered and in cases where exact optimal tests are not available we adopt Neyman's approach.

Chapter 4 is devoted to some optimal testing problems in mixtures where one or both of the component densities are in the form of infinite series. As an example, we consider the family of symmetric wrapped stable (SWS)- circular uniform (CU) mixtures having applications in the analysis of directional data. The family of SWS distributions (Mardia, 1972; SenGupta, 1998a) is now-a-days being widely used for modelling circular data. Mixtures of CU with the former have also applications as a larger family of circular distributions to incorporate possible outliers. We consider a mixture model with density (SenGupta and Pal, 2001)

$$g(\theta; a, \rho, \mu_0, p) = pf(\theta; a, \rho, \mu_0) + q(2\pi)^{-1}$$

where $f(\theta; a, \rho, \mu_0)$ is the density of a SWS distribution with parameters a , ρ and μ_0 . Restricting to such a mixture and using a variant of Wijsman's representation of the ratio of the densities of a maximal invariant under the alternative and the null hypotheses (Bhattacharyya and Johnson, 1969), we derive the LMPI test for the hypothesis of isotropy or randomness of directions — expressed in terms of the null value of the mixing proportion p in the model. Asymptotic distributions of the LMPI test statistic under both the null and the alternative hypotheses are also derived. Global monotonicity of the power function of the test is established exploiting some results on Fourier analysis from Wintner (1947). For the general situation when ρ is also unknown, the problem is tackled through the newly developed P^3 approach (SenGupta, 1991) and the solution is deferred till Chapter 7. The non-regular case of unknown a , the index parameter, is also discussed

and a Davies-motivated test (Davies, 1977, 1987) is proposed. Two real-life examples along with the associated rose diagrams, obtained by using the recently developed statistical package DDSTAP (SenGupta, 1998b), are presented.

In Chapter 5 we take up a multivariate model, *viz.*, the family of symmetric multivariate normal (SMN) mixtures, keeping in view its wide applicability in the statistical literature. We start with a p -mixture of k -variate standard SMN (SSMN) distributions with one component having the equicorrelation coefficient $\rho \in (-\frac{1}{k-1}, 1)$ and the other with $\rho = 0$, having the density

$$g_p(\mathbf{x}|m, \sigma^2, \rho) = p(2\pi\sigma^2)^{-k/2} \exp \left\{ -(\mathbf{x} - \mathbf{M})'(\mathbf{x} - \mathbf{M})/2\sigma^2 \right\} \\ + q(2\pi\sigma^2)^{-k/2} |\Sigma_\rho|^{-1/2} \exp \left\{ -(\mathbf{x} - \mathbf{M})'\Sigma_\rho^{-1}(\mathbf{x} - \mathbf{M})/2\sigma^2 \right\}$$

with $m = 0$ and $\sigma^2 = 1$ and attempt to construct some exact optimal tests. As the null hypothesis of ‘no contamination’ can be recast either as $\rho = 0$ or as $p = 1$, from theoretical and practical considerations based on statistical curvature, we advocate the test for the former assuming p to be known.

We next move on to the SMN mixture family with either or both of m and σ^2 appearing as the nuisance parameter(s). Three different cases are considered, *viz.*, m known, σ^2 unknown (Case 2); m, σ^2 both unknown (Case 3) and m unknown, σ^2 known (Case 4) (the case of both m and σ^2 known has been referred to as Case 1). Asymptotic distributions of the test statistics are derived. Some power tables are also presented to demonstrate the favourable performances of these tests.

The theory of $C(\alpha)$ tests developed by Neyman (1959) has been drawing attentions of great many researches for a long time. Construction of an optimal test in the class $C(\alpha)$ is based on certain Cramer-type regularity conditions, the existence of locally root n consistent estimators of the nuisance parameters present in the model under the null hypothesis H_0 and the theory of regression. In Chapter 6 we illustrate, with the help of a non-regular mixture of univariate normal distributions, that such an optimal test may not exist. It is therefore necessary to go for what we call ‘higher order’ $C(\alpha)$ tests. Under more stringent conditions we derive the general form of an optimal second order $C(\alpha)$ test and prove the main

theorem of this Chapter. It is noted that the theorem can be generalized to obtain optimal tests of third or higher orders, if necessary. The above theorem is then used to construct an optimal test for the hypothesis of ‘no contamination’ for the said example. Again, with reference to another example considered in Chapter 2, it is seen that even an optimal second order $C(\alpha)$ test may not exist. We construct an optimal third order test for this example and note that this coincides with the so called ‘skewness test’ (D’Agostino and Pearson, 1973, 1974; Mardia, 1974) for testing normality.

We next deal with the multivariate analogs of the univariate examples considered above and derive the union-intersection tests of Roy (1953) starting from the optimal tests obtained for the corresponding univariate models.

Chapter 7 deals with optimal testing problems for the hypothesis of ‘no mixture’ in two-component mixture distributions (with or without nuisance parameter(s)) in which the parameter of interest and the mixing proportion both are *unknown*. Attention is focussed on locally optimal procedures by considering what are called ‘L-optimal’ tests introduced by SenGupta (1991) and further pursued by Pal and SenGupta (2000a). The idea is to first recast the original several parametric hypotheses representation of the null hypothesis in terms of only a single hypothesis involving an appropriately chosen parametric function. Such a parametric function is termed as *Pivotal Parametric Product* (P^3 in short). After an appropriate P^3 is chosen, the test is constructed on the basis of a suitable (usually unbiased and/or consistent) estimator of it. We may call such a test a P^3 test. The (locally) optimal P^3 test, if exists and is unique, can then be seen to be the L-optimal test as defined by SenGupta (1991).

We begin by considering the case when there is no nuisance parameter and give a set of sufficient conditions on the class of densities such that the associated class of mixtures admits an appropriate P^3 . We also characterize a general form of the appropriate P^3 and the structure of the corresponding L-optimal test. Next we discuss the integrated likelihood approach and show that the LMP test for the parameter based on an integrated likelihood coincides with the L-optimal test. We discuss the applications of mixtures in reliability theory and note the existence of a rich class of mixtures of reliability distributions for which the forms of P^3 and the L-optimal test can

be conveniently derived by an appeal to the general result derived earlier. An asymptotic optimality property of the P^3 tests among the class of all tests based on linear unbiased estimating functions is also established.

Next we deal with the situation, with examples, where nuisance parameters are present in the model and extend the concept of L-optimal tests to that of L-optimal similar or L-optimal invariant or L-optimal $C(\alpha)$ tests. We also consider L-optimal tests in the SSMN and SMN mixture families of Chapter 5.

In Chapter 8 we take a different view-point for constructing asymptotically optimal tests in mixtures and attempt to propose tests for 'no contamination' which are optimal in terms of having the largest slope locally in the sense of Bahadur (1960; 1967). Such a test is called locally Bahadur optimal. We consider a two-component mixture of distributions with density

$$g(x|p, \theta) = pf(x|\theta) + (1 - p)f(x|\theta_0)$$

where $f(x|\theta)$ is the density (w.r.t. some σ -finite measure μ) of the one-parameter exponential family of distributions given by

$$f(x|\theta) = B(\theta) \exp [\psi(\theta)U(x)] h(x)$$

and propose a test for the parameter against two-sided alternatives, assuming the mixing proportion to be known. It may be noted that for such a testing problem, the LRT is not available in a closed form, neither its Bahadur optimality can be verified easily. We show that the proposed test is locally Bahadur optimal by verifying certain sufficient conditions of Hsieh (1979). We finally present some examples for each of which the conditions are seen to hold. As we note, the verification is non-trivial for all the examples, especially for the case of Poisson mixtures where we use some results on uniform convergence of series of functions.

Chapter 9 is devoted to optimal testing problems for multiparameter hypotheses in mixtures. Under a quite general setting, the locally most mean power unbiased (LMMPU) tests, developed by SenGupta and Vermeire (1986), lead many of the existing ones and can be derived by an appeal to the Neyman-Pearson theory. After presenting some relevant definitions and results from SenGupta and Vermeire (1986), we derive a general form of the LMMPU test for the hypothesis of 'no mixture' given by

$H_0 : \theta = (\theta_0, \theta_0, \dots, \theta_0)' = \theta_0 \mathbf{1}$ in a $(k+1)$ -component mixture of univariate uniparameter distributions having the joint density

$$g(x, \theta) = \prod_{j=1}^n [p_1 f(x_j, \theta_1) + p_2 f(x_j, \theta_2) + \dots + p_k f(x_j, \theta_k) + p_0 f(x_j, \theta_0)]$$

where $\theta = (\theta_1, \theta_2, \dots, \theta_k)$; $p_0 = 1 - p_1 - p_2 - \dots - p_k$; p_r 's are all known and θ_0 is also known. It is also assumed that the mixture distribution is identifiable. We next apply this result to the case of a mixture of normal distributions and derive the test exploiting the reflection principle as enunciated in Ferguson (1967, pp. 239-240). Verification of a condition regarding the positive definiteness of a matrix is still extremely difficult. However, for the particular case of $k = 2$, the above condition is seen to hold for small values of the level of significance, and hence the test so obtained is regular LMMPU. A two-component mixture of multivariate normal distributions is also considered.

We finally consider the case when the mixing proportion is unknown and take up the multivariate analogs of the univariate normal mixtures as considered in Chapter 7. Forms of the appropriate P^3 s and the corresponding union-intersection tests of Roy (1953) are derived.

In Chapter 10 we give an overview of the results obtained in some of the previous Chapters along with their significances and shortcomings and indicate some directions for further research.

Chapter 11 is the final Chapter of the thesis and is divided into three Sections. In Section 11.1 we present some selected computer programs, written in FORTRAN, relevant for Chapters 4 and 5. All the necessary tables are presented in Section 11.2. Section 11.3 is devoted to some diagrams obtained by using the statistical package DDSTAP.

After these eleven Chapters we present seven appendices A to G in which we reproduce some important definitions and results, with references, which have been used in different contexts throughout the thesis. The thesis is concluded with a list of references in the bibliography.

Chapter 2

Tests for Uniparameter Hypotheses in Regular Univariate Mixtures

2.1. Introduction

In this Chapter we consider the problem of optimal testing for uniparameter hypotheses in two-component mixtures of univariate distributions, with a brief review of the previous works wherever applicable. Since uniformly most powerful (UMP) tests do not exist in general, attention is restricted to LMP tests or different modifications/extensions of these to subsume nuisance parameters. The probability models we consider are ‘regular’ in the sense that usual LMP tests can be constructed on the basis of non-vanishing score statistics. This will also ensure that Efron’s statistical curvature (Efron, 1975), used for the purpose of initially evaluating the performance of an LMP test, can be defined in the usual way.

In Section 2.2 we briefly review the situation when there is no nuisance parameter present in the model. A brief discussion on statistical curvature is presented in Section 2.3. Section 2.4 deals with the case where nuisance parameters are present in the model. For location or scale or both location-scale families, the problems can be reduced by invariance, hence in most cases LMPI tests can be constructed. The approach of Spjøtvoll (1968) or Durairajan (1980) may also be applied to obtain LMPS or LMPU tests. All these tests are optimal in small samples. When no small sample optimal test exists one may appeal to the $C(\alpha)$ tests of Neyman (1959) or the tests of Davies (1977, 1987). We discuss these approaches in Subsection 2.4.1. In the next Subsection we give a general form of the LMPSU test following Spjøtvoll (1968). In Subsection 2.4.3 this result is applied to three examples of normal mixtures each reducible through invariance. Failure of this approach to any model leads us to restrict to the class of invariant tests and to derive the LMPIU test, if exists. As we note, one of the models

does not even admit unique LMPIU test, so we explore the possibility of applying Neyman's theory (1959) for this model. Part of this Chapter is based on Pal and SenGupta (2000b).

2.2. Nuisance parameters not present : LMP tests

2.2.1. Component distributions completely specified : test for no mixture

The model is

$$g(x|p) = pf_1(x) + qf_2(x) \quad (2.2.1)$$

where f_1 and f_2 are two completely specified densities w.r.t. some σ -finite measure and are not necessarily of the same family, p ($0 \leq p \leq 1$) is unknown and $q = 1 - p$. The hypothesis to be tested is $H_0 : p = 0$ against $H_1 : p > 0$ (or $H'_0 : p = 1$ against $H'_1 : p < 1$) on the basis of n i.i.d. observations $X_1, X_2, \dots, X_n$ drawn from (2.2.1). Under the usual assumption that the power function of any test is continuously differentiable under the integral sign w.r.t. p at the null value of the parameter, the form of the LMP (nonrandomized) test for testing H_0 against H_1 is given by (Durairajan and Kale, 1979; Durairajan, 1980):

$$\text{Reject } H_0 \text{ if } \sum_{i=1}^n \frac{f_1(X_i)}{f_2(X_i)} > c \quad (2.2.2)$$

where c is to be determined from the size condition. Usual reversal of the inequality in (2.2.2) holds if one considers testing H'_0 against H'_1 . Durairajan (1980) considered the asymptotic null distribution of the test statistic along with some examples on the above model.

2.2.2. One component density completely specified : test for no contamination

In this case the probability model is

$$g(x|p, \theta) = pf(x|\theta) + qf(x|\theta_0). \quad (2.2.3)$$

Here θ is an unknown parameter belonging to a parameter space $\Theta \subset \mathcal{R}$ and θ_0 is a known point of Θ . Such a model is known as a contamination model and is used to incorporate outliers in the data. The mixing proportion

p may be known or unknown. We want to test H_0 : 'no contamination' against the alternative H_1 : 'there is contamination'. Suppose p is known and lies in $(0, 1)$. H_0 is then equivalent to $H_0 : \theta = \theta_0$ which may be tested against one- or two-sided alternatives. Durairajan (1980) derived the LMP and LMPU tests for the two types of alternatives respectively. We present here the forms of the (non-randomized) tests in the following

Theorem 2.2.1. (a) If the power function of any test is continuously differentiable under the integral sign w.r.t. θ at θ_0 , for testing H_0 against the one-sided alternative $H_1 : \theta > (<) \theta_0$, the unique LMP test is of the form

$$\begin{aligned}\psi_0(\mathbf{x}) &= 1 & \text{if } \sum_{i=1}^n \left. \frac{\partial \log f(x_i|\theta)}{\partial \theta} \right|_{\theta_0} > (<) c \\ &= \gamma(\mathbf{x}) & \text{if } \sum_{i=1}^n \left. \frac{\partial \log f(x_i|\theta)}{\partial \theta} \right|_{\theta_0} = c \\ &= 0 & \text{if } \sum_{i=1}^n \left. \frac{\partial \log f(x_i|\theta)}{\partial \theta} \right|_{\theta_0} < (>) c\end{aligned}$$

where c and γ are to be determined from the size condition.

(b) If the power function of any test is continuously differentiable twice under the integral sign w.r.t. θ at θ_0 , for testing H_0 against the two-sided alternative $H : \theta \neq \theta_0$, the unique LMPU test has the form

$$\begin{aligned}\psi_0(\mathbf{x}) &= 1 & \text{if } T > k_1 + k_2 U \\ &= \gamma(\mathbf{x}) & \text{if } T = k_1 + k_2 U \\ &= 0 & \text{if } T < k_1 + k_2 U\end{aligned}$$

where

$$\begin{aligned}T &= p \left[\sum_{i=1}^n \left. \frac{\partial \log f(x_i|\theta)}{\partial \theta} \right|_{\theta_0} \right]^2 + q \sum_{i=1}^n \left[\left. \frac{\partial \log f(x_i|\theta)}{\partial \theta} \right|_{\theta_0} \right]^2 \\ &\quad + \sum_{i=1}^n \left. \frac{\partial^2 \log f(x_i|\theta)}{\partial \theta^2} \right|_{\theta_0}, \\ U &= \sum_{i=1}^n \left. \frac{\partial \log f(x_i|\theta)}{\partial \theta} \right|_{\theta_0}\end{aligned}$$

and k_1, k_2 and γ are to be determined from the size and local unbiasedness conditions.

Remark 2.2.1. Often it happens that k_2 above is zero, so that the LMPU test reduces to a simpler form. Also in many cases it is convenient to first construct a test without the local unbiasedness condition and then to verify this condition. If local unbiasedness is satisfied, the resulting test turns out to be optimal in a larger class of tests.

Next suppose p is unknown. If the value of θ of the contaminating density $f(x|\theta)$ happens to be known, the test for no contamination reduces to testing for the null value of p against the alternative that $p > 0$. In that case one can follow the results of Subsection 2.2.1.

2.3. Statistical curvature

For a one-parameter non-exponential family, Efron (1975, p. 1196) suggested a general definition of statistical curvature, γ_θ , as a measure to quantify how ‘nearly exponential’ these families are. Efron suggested that a value of $\gamma_{\theta_0}^2 < 1/8$ is not ‘large’ and one can expect linear methods to work ‘well’ in such a case. In the context of testing of hypotheses, in case a UMP test does not exist, LMP test may then be considered. Further in repeated sampling situations, the curvature ${}_n\gamma_{\theta_0}^2$ based on n i.i.d. observations, satisfies ${}_n\gamma_{\theta_0}^2 = \gamma_{\theta_0}^2/n$ and hence one can determine the sample size n which reduces the curvature below $1/8$. Let $\mathcal{F} = \{f_\theta(x), \theta \in \Theta\}$ be an arbitrary family of one-parameter density functions. Let $l_\theta(x) = \log f_\theta(x)$. Then, under the usual assumptions of the existence of the derivatives $\dot{l}_\theta$ and $\ddot{l}_\theta$, the statistical curvature of f at θ is defined as γ_θ given by

$$\gamma_\theta = \left(|\mathbf{M}_\theta| / i_\theta^3 \right)^{1/2} \quad \text{where} \quad \mathbf{M}_\theta \equiv \begin{pmatrix} E_\theta(\dot{l}_\theta^2) & E_\theta(\dot{l}_\theta \ddot{l}_\theta) \\ E_\theta(\dot{l}_\theta \ddot{l}_\theta) & E_\theta(\ddot{l}_\theta^2) - i_\theta^2 \end{pmatrix},$$

provided $i_\theta = E_\theta(\dot{l}_\theta^2) > 0$.

Since an identifiable mixture model usually constitutes a non-exponential family, this formula can be used to calculate the curvature of any one-parameter mixture family to judge the performance of the LMP test. This calculation is also particularly useful for the appropriate choice of the LMP test when the hypothesis of no contamination can be recast in more than one way in terms of different parameters occurring in the model. We shall consider such a situation in Chapter 5.

2.4. Nuisance parameters present in the model

The presence of nuisance parameters in a probability model is a matter of concern for any inference problem. In an exact (small sample) optimal testing procedure, nuisance parameters are generally got rid of by reducing the problem by any of the principles of sufficiency, similarity or invariance. It is also possible to construct a test based on the *integrated likelihood* (Aitkin and Rubin, 1985; Pal and SenGupta, 2000a) by considering a prior distribution for the nuisance parameter. We refer to Chapter 7 for a discussion on this approach. When no such reduction is possible one may appeal to the asymptotic procedures, *e.g.*, of Neyman (1959) or of Davies (1977, 1987) – the later being applicable when nuisance parameters are present only under the alternative.

2.4.1. Asymptotically optimal tests

Consider the density (2.2.3) where both θ and p are unknown and $0 < p < 1$ and $\theta \geq \theta_0$. Clearly either of θ and p can be considered as a nuisance parameter when testing for the other. Suppose we want to test the null hypothesis $H_0 : \theta = \theta_0$. Since $g(x|p, \theta)$ is linear in p (Titterington *et al.*, 1985, p. 158) the locally optimal score statistic $Z_n(p)$ of Neyman (1959) is independent of p and is given by

$$Z_n(p) = Z_n = (\sqrt{n}\gamma)^{-1} \sum_{i=1}^n \frac{\partial}{\partial \theta} \log f(X_i|\theta) \Big|_{\theta_0}$$

with

$$\gamma^2 = \text{Var}_{\theta_0} \left[\frac{\partial}{\partial \theta} \log f(X|\theta) \Big|_{\theta_0} \right].$$

Now interchange the roles of θ and p and consider the problem of testing $H_0 : p = 0$ when θ lies in a closed interval, say $[L, U]$. Evaluate the corresponding normalized score function $Z_n(\theta)$ following Davies (1977). This score function is not independent of θ and so ‘a straightforward critical region cannot be constructed’ (Titterington *et al.*, 1985, p. 158). A conservative procedure proposed by Davies (1977) is based on a critical region

$$\sup_{\theta \in [L, U]} Z_n(\theta) > c.$$

Davies (1977) discussed how to calculate an upper bound of the significance level (p-value) of this test using the theory of total variation. The test has also the optimum property in the sense that its power for any fixed θ tends to that of the locally optimal test as the significance level (α) tends to 0.

We consider some examples in Chapter 3 where these approaches have been adopted.

2.4.2. Exact optimal tests

The general theory on the existence and construction of exact optimal tests such as LMPU, LMPI and LMPS tests for mixtures has been discussed by Durairajan (1980) and Durairajan and Kale (1982). Durairajan (1980) derived LMPU tests for the parameter of interest against two-sided alternatives when no nuisance parameter is present in the mixture model. However the presence of nuisance parameters is also of practical importance and, in some cases, can be tackled by the approach of Spjøtvoll (1968). We present below a general result on the form of the LMPSU test in a two-component mixture.

Let

$$g(x|p, \vartheta, \theta) = pf(x|\vartheta, \theta) + qf(x|\vartheta, \theta_0) \quad (2.4.1)$$

be a mixture density in which p ($0 < p < 1$) is known, $q = 1 - p$, θ is the unknown parameter of interest in $\Theta \subset \mathcal{R}$, θ_0 is a known point of Θ , ϑ is the unknown nuisance parameter (may be vector-valued) and f is a density with respect to some σ -finite measure of a real-valued random variable. A random sample $(X_1, \dots, X_n)$ of size n is drawn from (2.4.1). We want to test $H_0 : \theta = \theta_0$ against the two-sided alternative $H : \theta \neq \theta_0$. We assume that the family of the distributions of $(X_1, \dots, X_n)$ under H_0 admits a boundedly complete sufficient statistic, say T .

Theorem 2.4.1. Suppose that for any test ψ the first and second derivative with respect to θ of the power function $\beta_\psi(\theta, \vartheta)$ can be computed under the integral sign. Then among all similar level α tests unbiased at θ_0 the following test ψ_0 maximizes the second derivative of the power function at (θ_0, ϑ_0)

$$\begin{aligned} \psi_0(\mathbf{x}) &= 1 && \text{if } V > k_1(t) + k_2(t) U \\ &= \gamma(t) && \text{if } V = k_1(t) + k_2(t) U \\ &= 0 && \text{if } V < k_1(t) + k_2(t) U \end{aligned}$$

where

$$V = p \left[\sum_{i=1}^n \frac{\partial \log f(x_i | \vartheta_0, \theta)}{\partial \theta_0} \right]^2 + q \sum_{i=1}^n \left[\frac{\partial \log f(x_i | \vartheta_0, \theta)}{\partial \theta_0} \right]^2 + \sum_{i=1}^n \frac{\partial^2 \log f(x_i | \vartheta_0, \theta)}{\partial \theta_0^2},$$

$$U = \sum_{i=1}^n \frac{\partial \log f(x_i | \vartheta_0, \theta)}{\partial \theta_0},$$

$k_1(t)$, $k_2(t)$ and $\gamma(t)$ are constants depending on a fixed value t of T and are to be determined from $E_{\theta_0}^{X|t} \psi_0(X) = \alpha$ and $\left. \frac{\partial}{\partial \theta} E_{\theta}^{X|t} \psi_0(X) \right|_{\theta_0} = 0$, i.e., the size and local unbiasedness conditions respectively. Further if ψ_0 does not depend on ϑ_0 and is unique, then it is the LMPSU test for H_0 against H .

Proof. If $L(x|p, \vartheta, \theta) = \prod_{i=1}^n \{p f(x_i | \vartheta, \theta) + q f(x_i | \vartheta_0, \theta)\}$ denotes the joint density at $x = (x_1, \dots, x_n)$, it can be verified that

$$\frac{\partial \log L(x|p, \vartheta_0, \theta)}{\partial \theta_0} = p \sum_{i=1}^n \frac{\partial \log f(x_i | \vartheta_0, \theta)}{\partial \theta_0}$$

and

$$\frac{\partial^2 \log L(x|p, \vartheta_0, \theta)}{\partial \theta_0^2} = pq \sum_{i=1}^n \left[\frac{\partial \log f(x_i | \vartheta_0, \theta)}{\partial \theta_0} \right]^2 + p \sum_{i=1}^n \frac{\partial^2 \log f(x_i | \vartheta_0, \theta)}{\partial \theta_0^2}.$$

The theorem now follows in the lines of Theorem 3.3 of Spjøtvoll (1968).

□

Remark 2.4.1. The above theorem allows us to find the optimal test among the class of similar tests having the restriction of local unbiasedness. As in the case of no nuisance parameter, it is also sometimes convenient to first construct a test without the local unbiasedness condition and then to verify whether this condition holds.

2.4.3. Optimal tests in univariate normal mixtures

Mixtures of normal distributions have been used extensively as models in a wide variety of important practical situations. Examples of such situations include, among others, crab forehead sizes (Pearson, 1894), height differences in plant pairs (Box and Tiao, 1968), blood and urine glucose in Seminole Indians (Elston *et al.*, 1974), remote sensors on orbiting satellites (Odell, 1976), discriminating pressure in dental experiments (Marks and Rao, 1979) and so on. There are also several tests for the detection of two-component mixtures of normals (see Mendell *et al.*, 1993 and the references therein). But very few of them possess small sample (exact) optimality properties based on Neyman-Pearson theory or asymptotic properties such as optimal $C(\alpha)$ of Neyman (1959). We shall consider the following three models of univariate normal mixtures and apply Theorem 2.4.1 on each to obtain the LMPSU test, if exists.

- (i) $g(x|p, \mu_1, \mu_2) = p\phi(x - \mu_1) + q\phi(x - \mu_2)$,
- (ii) $g(x|p, \mu, \tau^2) = \frac{p}{\tau}\phi\left(\frac{x - \mu}{\tau}\right) + \frac{q}{\tau}\phi\left(\frac{x}{\tau}\right)$ and
- (iii) $g(x|p, \mu_1, \mu_2, \tau^2) = \frac{p}{\tau}\phi\left(\frac{x - \mu_1}{\tau}\right) + \frac{q}{\tau}\phi\left(\frac{x - \mu_2}{\tau}\right)$,

where in each of the models $\phi(\cdot)$ is the density of a standard normal variable, p is the known mixing proportion, $q = 1 - p$ and all other parameters are unspecified. Making an appropriate reparametrization, each of the models can be expressed in an equivalent form, given respectively, by

- (i)' $g(x|p, \mu, \theta) = p\phi(x - \mu - \theta) + q\phi(x - \mu)$,
- (ii)' $g(x|p, \sigma^2, \theta) = \frac{p}{\sigma}\phi\left(\frac{x}{\sigma} - \theta\right) + \frac{q}{\sigma}\phi\left(\frac{x}{\sigma}\right)$ and
- (iii)' $g(x|p, \mu, \sigma^2, \theta) = \frac{p}{\sigma}\phi\left(\frac{x - \mu}{\sigma} - \theta\right) + \frac{q}{\sigma}\phi\left(\frac{x - \mu}{\sigma}\right)$.

Clearly (i)', (ii)' and (iii)' are reducible through invariance with respect to changes of location, scale and location-scale respectively. The model (i)' appears in Marks and Rao (1979) in connection with estimation of discriminating pressure in dental experiments. Durairajan (1980) considered the problem of constructing LMPS and LMPI tests for parameters in two-component mixtures of normal distributions, as examples, in the case of one-sided alternatives. We consider here testing for 'no mixture' against

two-sided alternatives because these are more reasonable for the particular models chosen. Also, as we shall see later, the resulting univariate tests can be combined appropriately to obtain reasonable tests for the corresponding multivariate situations.

Theorem 2.4.2. For testing $H_0 : \theta = 0$ against $H_1 : \theta \neq 0$, the unconditional LMPSU test exists for the model (i') and consists in rejecting H_0 for large values of $\sum_{i=1}^n (X_i - \bar{X})^2$. For models (ii') and (iii'), no unconditional LMPSU test exists.

Proof. Let us denote the statistics V and U of Theorem 2.4.1 by (V_1, U_1) , (V_2, U_2) and (V_3, U_3) for (i)', (ii)' and (iii)' respectively. It is easy to see that

$$\begin{aligned} V_3 &= p \left[\sum_{i=1}^n \left(\frac{x_i - \mu_0}{\sigma_0} \right) \right]^2 + q \sum_{i=1}^n \left(\frac{x_i - \mu_0}{\sigma_0} \right)^2 - n \\ \text{and } U_3 &= \sum_{i=1}^n \left(\frac{x_i - \mu_0}{\sigma_0} \right). \end{aligned} \quad (2.4.2)$$

The corresponding expressions for (V_1, U_1) and (V_2, U_2) are obtained from (2.4.2) by putting $\sigma_0 = 1$ and $\mu_0 = 0$ respectively. Boundedly complete sufficient statistics under the null hypotheses in the respective cases are $\bar{X}$, $\sum_i X_i^2$ and $(\bar{X}, \sum_i (X_i - \bar{X})^2)$. Also verification of the conditions of Theorem 2.4.1 for each of the cases can be accomplished by first expanding the finite product in the mixture likelihood under the integral sign appearing in the expression for the power function of any test and then applying Theorem 9 of Lehmann (1986, p. 59) (see Result G1 of Appendix G).

After a straightforward derivation following standard arguments through Basu's theorem it turns out that the unconditional LMPSU test in the first case is

$$\begin{aligned} \psi_0^{(1)}(x) &= 1 \quad \text{if } S^2 = \sum_i (x_i - \bar{x})^2 > c \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

For each of the last two cases, it turns out that Basu's theorem is not applicable and hence no unconditional LMPSU test exists. $\square$

Remark 2.4.2. Observe that the test $\psi_0^{(1)}(x)$ is invariant with respect to change of location and is, therefore, LMPIU. Further, this test is equivalent to the χ^2 test and hence the cut-off points can be found quite easily.

The non-existence of the unconditional LMPSU tests for the last two models naturally leads us to derive the LMPIU tests for these problems. Consider first (ii)'. A maximal invariant for this model is

$$Y = (Y_1, \dots, Y_{n-1})$$

where $Y_i = X_i/X_n$, $i = 1, \dots, n-1$. The density of Y is

$$g_Y(y|\theta) = \int_{-\infty}^{\infty} \prod_{i=1}^n \{p\phi(y_i w - \theta) + q\phi(y_i w)\} |w|^{n-1} dw \quad (2.4.3)$$

with $y_n \equiv 1$. Let $\psi = \psi(y)$ be any invariant test with power function $\beta_\psi(\theta)$. Using again the expansion of the finite product in the mixture likelihood, one can also write

$$g_Y(y|\theta) = \text{Const.} \sum_{r=0}^n \sum_{C_r} p^r q^{n-r} \exp[-r\theta^2/2] \\ \cdot \int_{-\infty}^{\infty} \exp\left\{\theta w \sum_{i \in I_r} y_i\right\} \exp\left\{-w^2 \left(\sum_{i=1}^n y_i^2\right) / 2\right\} |w|^{n-1} dw,$$

where for $r = 1, 2, \dots, n$, I_r denotes a typical set of r integers of the form $(i_1, i_2, \dots, i_r)$ out of $(1, 2, \dots, n)$, I_0 is the null set and C_r is the collection of all possible sets I_r . By Theorem 9 of Lehmann (1986, p. 59) the above integral can be differentiated any number of times w.r.t. θ and the derivatives can be passed under the sign of integration. Hence the right hand side of (2.4.3) is also differentiable any number of times and the derivatives can be passed under the integral sign. The same is also true for the power function

$$\beta_\psi(\theta) = \text{Const.} \sum_{r=0}^n \sum_{C_r} p^r q^{n-r} \exp[-r\theta^2/2] \\ \cdot \int_{R^n} \psi(y) \exp\left[\theta w \sum_{i \in I_r} y_i\right] \times \\ \exp\left[-w^2 \left(\sum_{i=1}^n y_i^2\right) / 2\right] |w|^{n-1} dw dy_1 \dots dy_{n-1}.$$

One can easily verify that

$$\left. \frac{\partial}{\partial \theta} g_Y(y|\theta) \right|_0 = 0.$$

The form of the LMPIU test for testing H_0 against H then reduces to

$$\begin{aligned}\psi_0^{(2)}(y) &= 1 && \text{if } \left. \frac{\partial^2}{\partial \theta^2} g_Y(y|\theta) \right|_0 > k g_Y(y|0) \\ &= 0 && \text{otherwise.}\end{aligned}$$

Simplification yields

$$\begin{aligned}g_Y(y|0) &= \text{Const.} \int_{-\infty}^{\infty} \exp \left\{ -\frac{n(\bar{y}^2 + s_y^2)w^2}{2} \right\} |w|^{n-1} dw \\ &= \text{Const.} (\bar{y}^2 + s_y^2)^{-n/2},\end{aligned}$$

where $\bar{y} = \sum_{i=1}^n y_i/n$ and $s_y^2 = \sum_{i=1}^n (y_i - \bar{y})^2/n$.

$\left. \frac{\partial^2}{\partial \theta^2} g_Y(y|\theta) \right|_0$ needs some more algebra and simplifies to

$$\text{Const.} (k_1 \bar{y}^2 + k_2 s_y^2)(\bar{y}^2 + s_y^2)^{-(n+2)/2} - k_3 (\bar{y}^2 + s_y^2)^{-n/2},$$

with $k_1 = ng$, $k_2 = n^2p + ng$ and Const. representing a positive constant. Hence noting that the ratio $(k_1 \bar{y}^2 + k_2 s_y^2)/(\bar{y}^2 + s_y^2)$ is decreasing in $\bar{y}^2/s_y^2$, the test $\psi_0^{(2)}$ reduces to

$$\begin{aligned}\psi_0^{(2)}(y) &= 1 && \text{if } \bar{y}^2/s_y^2 < c \\ &= 0 && \text{otherwise.}\end{aligned}$$

In terms of the original variables $X_1, \dots, X_n$, this test is

$$\begin{aligned}\psi_0^{(2)}(x) &= 1 && \text{if } \bar{x}^2/s_x^2 < c \\ &= 0 && \text{otherwise.}\end{aligned}$$

This test is obviously equivalent to the t -test and hence the cut-off points can be easily found.

Finally we consider (iii)'. Here a maximal invariant is

$$Z = (Z_2, \dots, Z_{n-1}),$$

where $Z_i = (X_i - X_1)/(X_n - X_1)$, $i = 2, \dots, n-1$. The density of Z is

$$g_Z(z|\theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{i=1}^n \{p\phi(z_i w_2 + w_1 - \theta) + q\phi(z_i w_2 + w_1)\} |w_2|^{n-2} dw_1 dw_2$$

with $z_1 \equiv 0$ and $z_n \equiv 1$. Verification of the conditions again follows as before. Also it can be shown that

$$\left. \frac{\partial}{\partial \theta} g_Z(z|\theta) \right|_0 = 0.$$

Hence, in terms of Z , the LMPIU test for testing $H_0 : \theta = 0$ against $H : \theta \neq 0$ is given by

$$\begin{aligned} \psi_0^{(3)}(z) &= 1 \quad \text{if } \frac{\left. \frac{\partial^2}{\partial \theta^2} g_Z(z|\theta) \right|_0}{g_Z(z|0)} > k \\ &= 0 \quad \text{otherwise.} \end{aligned} \quad (2.4.4)$$

First consider the denominator $g_Z(z|0)$ in (2.4.4). This is equal to

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{i=1}^n \phi(z_i w_2 + w_1) |w_2|^{n-2} dw_1 dw_2.$$

Note that the term $\prod_{i=1}^n \phi(z_i w_2 + w_1)$ can be written as

$$\frac{k}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left[-\frac{1}{2(1-\rho^2)} \left\{ \frac{w_2^2}{\sigma_2^2} - \frac{2\rho w_1 w_2}{\sigma_1\sigma_2} + \frac{w_1^2}{\sigma_1^2} \right\} \right],$$

where

$$\rho = -\frac{b}{\sqrt{ca}}, \quad \sigma_1^2 = \frac{a}{ac-b^2}, \quad \sigma_2^2 = \frac{c}{ac-b^2} \quad \text{and } k = \text{Const.} \sigma_1 \sigma_2 \sqrt{1-\rho^2}$$

with

$$a = \sum_{i=1}^n z_i^2, \quad b = \sum_{i=1}^n z_i \quad \text{and } c = n.$$

Hence

$$\begin{aligned} g_Z(z|0) &= \text{Const.} \sigma_1 \sigma_2 \sqrt{1-\rho^2} E \left\{ |W_2|^{n-2} \mid W_2 \sim N(0, \sigma_2^2) \right\} \\ &= \text{Const.} (s_z^2)^{-(n-1)/2}, \end{aligned}$$

where $s_z^2 = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2$.

Now consider the numerator. We have

$$\begin{aligned} & \left. \frac{\partial^2}{\partial \theta^2} \prod_{i=1}^n \{p\phi(z_i w_2 + w_1 - \theta) + q\phi(z_i w_2 + w_1)\} \right|_{\theta=0} \\ &= \left\{ \lambda_1 w_1^2 + 2\lambda_{12} w_1 w_2 + \lambda_2 w_2^2 + np \right\} \prod_{i=1}^n \phi(z_i w_2 + w_1) \end{aligned}$$

after some simplification, where

$$\lambda_1 = n^2 p^2 + npq, \quad \lambda_2 = n^2 p^2 \bar{z}^2 + pq \sum_{i=1}^n z_i^2 \quad \text{and} \quad \lambda_{12} = n(np^2 + pq)\bar{z}.$$

Hence

$$\begin{aligned} & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left. \frac{\partial^2}{\partial \theta^2} \prod_{i=1}^n \{p\phi(z_i w_2 + w_1 - \theta) + q\phi(z_i w_2 + w_1)\} \right|_{\theta=0} |w_2|^{n-2} dw_1 dw_2 \\ &= \text{Const.} \quad \sigma_1 \sigma_2 \sqrt{1 - \rho^2} E \left[|W_2|^{n-2} \left\{ \lambda_1 W_1^2 + 2\lambda_{12} W_1 W_2 + \lambda_2 W_2^2 + np \right\} \right] \end{aligned}$$

where $(W_1, W_2) \sim \text{BN}(0, 0, \sigma_1^2, \sigma_2^2, \rho)$.

Writing the above expectation $E[\cdot]$ as $E_{W_2} E_{W_1|W_2}[\cdot]$ and making some algebraic manipulations, the above double integral simplifies to

$$k_1 (s_z^2)^{-(n-1)/2} + k_2 \bar{z}^2 (s_z^2)^{-(n+1)/2} \quad \text{for some constants } k_1 \text{ and } k_2 > 0.$$

The test then reduces to

$$\begin{aligned} \psi_0^{(3)}(z) &= 1 \quad \text{if } \frac{\bar{z}^2}{s_z^2} > c \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

In terms of the original variables $X_1, \dots, X_n$, the test is

$$\begin{aligned} \psi_0^{(3)}(x) &= 1 \quad \text{if } \frac{(x_1 - \bar{x})^2}{s_x^2} > c \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

The null distribution of $\frac{X_1 - \bar{X}}{\sqrt{n}s_X}$ is known (see, *e.g.*, Lehmann and Casella, 1998, p. 94) and the cut-off points can be found from this distribution by numerical techniques. This test, however, is not unique as it depends on the particular (e.g. in terms of X_1) maximal invariant chosen. So no unique LMPIU test for the model (iii)' exists. Also the expectation of $(X_1 - \bar{X})^2/s_X^2$ under the alternative hypotheses is independent of θ . Hence the test does not even qualify as a P^3 test (see Chapter 7). For these reasons one may explore the plausibility of applying Neyman's $C(\alpha)$ test for this problem. We shall consider these in detail in Chapter 6.

Chapter 3

Tests for Uniparameter Hypotheses in Non-regular Univariate Mixtures

3.1. Introduction

The probability models considered in Chapter 2 are all 'regular' in the sense that for each of these models the score statistic $\partial \log g / \partial \theta|_{\theta_0}$ does not vanish identically, so that this can be used to construct the LMP test for the hypothesis $H_0 : \theta = \theta_0$. There are, however, models for which the score statistic vanishes identically in x , hence the usual LMP test cannot be constructed. Also Efron's statistical curvature (Efron, 1975), used to impart an initial evaluation of an LMP test in non-exponential families, is not defined in the usual way. Moreover there are models in which nuisance parameters are present only under the alternative hypotheses, so that Neyman's $C(\alpha)$ test procedure (Neyman, 1959) cannot be applied. We call such models 'non-regular'. We note that the first kind of non-regularity can be taken care of by a suitable reparametrization (Cox and Hinkley, 1974, pp. 117-118), whereas those of the second kind can be tackled by the approach of Davies (1977, 1987).

In this Chapter we consider some two-component mixtures of univariate normal distributions depicting such non-regularities and construct optimal tests for the parameter or the mixing proportion involved in the models. We propose a general normal mixture model in Section 3.2 and explore the constructions of optimal tests of the parameter of interest under a variety of situations regarding the knowledge of the other parameter or the mixing proportion. In Section 3.3 we consider the problem of testing for the mixing proportion p . We note that, when nuisance parameters are present, LMPS tests do not exist. We derive optimal $C(\alpha)$ tests in these situations.

3.2. Tests for the parameter

The general model proposed here is

$$g(x|p, \mu, \sigma^2) = pN(x|\mu, \sigma^2) + (1 - p)N(x|-\mu, \sigma^2) \quad (3.2.1)$$

where $N(x|\mu, \sigma^2)$ is the p.d.f. of a normal distribution with mean μ and variance σ^2 . Note that $0 \leq p \leq 1$, $\sigma^2 > 0$ and μ is assumed to belong to $[0, \infty)$ to ensure identifiability. We first consider the problem of testing $H_0 : \mu = 0$ against $H_1 : \mu > 0$ and construct LMP or $C(\alpha)$ tests under different situations concerning knowledge of the other parameters p and σ^2 . The role of Efron's statistical curvature (Efron, 1975) in imparting an initial evaluation of an LMP test for use in practice in this non-exponential family is also exploited.

Case I $p = 1/2$: In this case σ^2 may be known, equal to 1, say, or unknown.

(a) $\sigma^2 = 1$: (3.2.1) then becomes

$$g(x|\mu) = (1/2) [\phi(x - \mu) + \phi(x + \mu)] \quad (3.2.2)$$

where $\phi(\cdot)$ is the p.d.f. of a standard normal variable. A model of this type is known as mirror-image model. Here $\mu = 0$ is a *singular point* in the sense that any unbiased estimator of μ based on a single observation from (3.2.2), if exists, has infinite variance at $\mu = 0$ (see Lehmann and Casella, 1998, p. 139). Bartels (1984) considered the $\frac{1}{2}$ -mixture of circular normal distributions with location parameters of opposite signs, to model the movement of certain biological organisms. Our model (3.2.2) (possibly even in the presence of unknown σ^2) provides an obvious extension from applications in directional data to quite general situations. Another very closely related model to (3.2.2) is the folded normal distribution given by the p.d.f.

$$h(x|\mu) = \phi(x - \mu) + \phi(x + \mu) \quad (3.2.3)$$

with $x \geq 0, \mu \geq 0$ (see Johnson and Leone, 1977, p. 153), which can also be viewed as a $\frac{1}{2}$ -mixture of two folded normal distributions with means μ and $-\mu$ and each folded at '0'.

Note that for both the models (3.2.2) and (3.2.3), the score function $\partial \log g(x|\mu)/\partial \mu|_{\mu=0}$ or $\partial \log h(x|\mu)/\partial \mu|_{\mu=0}$ vanishes identically in x , so that usual LMP test for H_0 can not be found. Moreover, Efron's curvature is not defined in the usual way. However, a simple unified reparametrization is possible for both the models under which regular LMP tests can be constructed whose performance can be judged through Efron's curvature. To see this, observe that (3.2.2) can be rewritten as (SenGupta and Pal, 1993a)

$$g(x|\theta) = (2\pi)^{-1/2} \exp \left\{ -(x^2 + \theta)/2 \right\} \sum_{\alpha=0}^{\infty} \theta^{\alpha} x^{2\alpha} / (2\alpha)! ; \quad \theta = \mu^2 \geq 0.$$

Similar reparametrization goes through for model (3.2.3) also.

The LMP test: Note that H_0 and H_1 can be equivalently recast as $H_0 : \theta = 0$ and $H_1 : \theta > 0$ respectively. Now $\partial \log g(x|\theta)/\partial \theta|_{\theta=0} = (x^2 - 1)/2$ for the reparametrized forms of both (3.2.2) and (3.2.3). So, based on n i.i.d. observations $X_1, X_2, \dots, X_n$, the LMP test rejects H_0 in favour of H_1 if $\sum_{i=1}^n X_i^2 > C$, where C is so chosen that the size of test has a specified value. Since under H_0 each of (3.2.2) and (3.2.3) reduces to a simple form, C can be easily computed.

Monotonicity of the power function: Consider model (3.2.2). Since the transformation $\theta = \mu^2$ from $[0, \infty)$ to $[0, \infty)$ is one to one, going back to original parametrization, one can easily verify that

$$\partial P [X^2 > C | \mu] / \partial \mu = \phi(C^{1/2} - \mu) - \phi(C^{1/2} + \mu) > 0 \quad \text{if } \mu > 0.$$

Thus $P [X^2 > C | \mu] \uparrow \mu > 0$. An application of Lemma 3.1 of SenGupta and Chang (2000) (see Result C2 of Appendix C) then ensures the monotonicity of the power function of the LMP test based on n i.i.d. observations. One can proceed in a similar way to establish monotonicity for model (3.2.3) also.

Consistency of the LMP test: Considering again model (3.2.2) and using Central Limit Theorem (CLT) one gets asymptotically

$$P \left[\sum_{i=1}^n X_i^2 > C \mid \mu \right] = 1 - \Phi \left[(v(\mu))^{-1} \left\{ n^{1/2} C^* - n^{1/2} (1 + \mu^2) \right\} \right]$$

$$\rightarrow 1 \text{ as } n \rightarrow \infty$$

where $v^2(\mu) = \text{Var}_\mu(X^2)$ and C^* corresponds to the asymptotic size of the test. This shows that the LMP test is consistent. Similar argument goes through in case of model (3.2.3) also.

Statistical curvature: For a one-parameter non-exponential family, Efron (1975, p. 1196) suggested a general definition of statistical curvature, γ_θ , as a measure to quantify how 'nearly exponential' these families are. We have already discussed this aspect in Chapter 2. Here we consider model (3.2.3) in reparametrized form and proceed to compute the value of the statistical curvature at $\theta = 0$. It is easy to verify for this model that with usual notations

$$i_\theta = (-1/2) + \left\{ \sum_{\alpha=0}^{\infty} \theta^\alpha x^{2\alpha} / (2\alpha)! \right\}^{-1} \sum_{\alpha=1}^{\infty} \alpha \theta^{\alpha-1} x^{2\alpha} / (2\alpha)!$$

and

$$\begin{aligned} \ddot{l}_\theta = & \left\{ \sum_{\alpha=0}^{\infty} \theta^\alpha x^{2\alpha} / (2\alpha)! \right\}^{-2} \left[\left\{ \sum_{\alpha=0}^{\infty} \theta^\alpha x^{2\alpha} / (2\alpha)! \right\} \right. \\ & \left. \cdot \left\{ \sum_{\alpha=2}^{\infty} \alpha(\alpha-1) \theta^{\alpha-2} x^{2\alpha} / (2\alpha)! \right\} - \left\{ \sum_{\alpha=1}^{\infty} \alpha \theta^{\alpha-1} x^{2\alpha} / (2\alpha)! \right\}^2 \right]. \end{aligned}$$

So $i_0 = (x^2 - 1)/2$ and $\ddot{l}_0 = -x^4/6$. Under $H_0 : \theta = 0$, the distribution reduces to that of a folded normal variable with p.d.f.

$$g(x|0) = (2/\pi)^{1/2} \exp(-x^2/2); x \geq 0.$$

Hence using the above expressions for i_0 and $\ddot{l}_0$ and noting that

$$(2/\pi)^{1/2} \int_0^\infty x^r \exp(-x^2/2) dx = \pi^{-1/2} 2^{r/2} \Gamma\left(\frac{r+1}{2}\right)$$

for positive integral values of r , it is straightforward to see that

$$i_0 = E_0(i_0^2) = 1/2, E_0(i_0 \ddot{l}_0) = -1 \text{ and } E_0(\ddot{l}_0^2) = 35/12.$$

Therefore

$$\mathbf{M}_0 = \begin{pmatrix} E_0(\dot{l}_0^2) & E_0(\dot{l}_0 \ddot{l}_0) \\ E_0(\dot{l}_0 \ddot{l}_0) & E_0(\ddot{l}_0^2) - \dot{l}_0^2 \end{pmatrix} = \begin{pmatrix} 1/2 & -1 \\ -1 & 8/3 \end{pmatrix}$$

and ${}_1\gamma_0^2 = |\mathbf{M}_0|/\dot{l}_0^3 = 8/3$. Since ${}_n\gamma_0^2 = {}_1\gamma_0^2/n$, it is then necessary to take a sample of size n greater than 21 to reduce the value of curvature below the 'worrisome point' $1/8$. The LMP test based on such a sample size can be expected to 'work well', *i.e.*, has a reasonable power performance.

(b) σ^2 unknown: Here also we come across a mirror-image model with p.d.f.

$$g(x|\mu, \sigma^2) = (1/2) [\sigma^{-1}\phi((x - \mu)/\sigma) + \sigma^{-1}\phi((x + \mu)/\sigma)] \quad (3.2.4)$$

with $-\infty < x < \infty, \mu \geq 0, \sigma^2 > 0$ or a closely related folded normal model

$$h(x|\mu, \sigma^2) = (1/2) [\sigma^{-1}\phi((x - \mu)/\sigma) + \sigma^{-1}\phi((x + \mu)/\sigma)] \quad (3.2.5)$$

with $x \geq 0, \mu \geq 0, \sigma^2 > 0$. Considering, *e.g.*, (3.2.4) in the reparametrized form with $\theta = \mu^2$, one can easily see that

$$\partial \log g(X|\theta) / \partial \theta|_{(\theta=0, \sigma^2=\sigma_0^2)} = (X^2 - \sigma_0^2) / 2\sigma_0^4$$

where σ_0^2 is a fixed value of the nuisance parameter σ^2 . Also $T = \sum_{i=1}^n X_i^2$ is a complete sufficient statistic for σ^2 under $H_0 : \theta = 0$. It is at once clear that the approach of getting an LMPS test for H_0 fails because such a test then essentially comes out in the form : Reject H_0 if $T > C(t)$ where t is the fixed value of the statistic T under H_0 . Similar argument holds for model (3.2.5) also.

In order to construct an optimal $C(\alpha)$ test for this problem we can proceed in the line of Neyman (1959). But we see (Chapter 6) that such a test does not exist in this case and we need, therefore, to go for higher order $C(\alpha)$ tests. Details are discussed in Chapter 6.

Case II p known and belongs to $(0, 1) \setminus \{1/2\}$: Two cases may be considered.

(a) $\sigma^2 = 1$: In this case simple LMP test can be found and it comes out as : Reject H_0 if $\bar{X} > (<) C$ when $p > (<) 1/2$. The cut-off point C can be easily computed since under H_0 , $X_1, X_2, \dots, X_n$ are i.i.d. $N(0, 1)$.

(b) σ^2 unknown: If p_0 is the known value of p one can easily see that

$$\partial \log g / \partial \mu|_{(\mu=0, \sigma^2=\sigma_0^2)} = (2p_0 - 1)x / \sigma_0^2.$$

Also $S = \sum_i X_i^2$ is a complete sufficient statistic for σ^2 under H_0 . The LMP similar test for H_0 then becomes : Reject H_0 if

$$\left(\sum_i X_i \right) / \left(\sum_i X_i^2 \right)^{1/2} > (<) C \text{ when } p_0 > (<) 1/2,$$

or equivalently if $t > t_\alpha (< t_{1-\alpha})$ when $p_0 > (<) 1/2$, where t has a central t distribution with $n - 1$ d.f. under H_0 .

Case III p unknown: Two situations may arise:

(a) $\sigma^2 = 1$: The model (3.2.1) then becomes

$$g(x|p, \mu) = pN(x|\mu, 1) + (1 - p)N(x|-\mu, 1).$$

Note that under $H_0 : \mu = 0$, p is not present in the model and occurs only under the alternative. So Davies' technique (1977) seems to be applicable in this case. For each fixed p_0 we start with an asymptotically optimal test for H_0 , which in our case, may be taken as the LMP test and then take supremum of a Gaussian process. We have, following the same notations of Davies,

$$\partial \log g / \partial \mu|_{\mu=0} = (2p_0 - 1)X$$

and

$$\{\gamma(p_0)\}^2 = \text{Var}_{H_0} [\partial \log g / \partial \mu|_{\mu=0}] = (2p_0 - 1)^2.$$

Also

$$Z_n(p_0) = n^{-1/2} |2p_0 - 1|^{-1} \sum_i (2p_0 - 1)X_i = \text{Sgn}(2p_0 - 1) n^{-1/2} \sum_i X_i.$$

As regards the covariance function $\rho(p_1, p_2) = \text{Cov}(Z_n(p_1), Z_n(p_2))$ under H_0 , it is easy to verify that

$$\rho(p_1, p_2) = 1 \text{ if both } p_1, p_2 > \text{ or } < 1/2 \text{ and } = -1 \text{ otherwise.}$$

Several consequences now creep in:

- (i) If $p \in (0, 1) \setminus \{1/2\}$, the function $\rho_{11}(p) = \frac{\partial^2}{\partial p_1^2} \rho(p_1, p) \Big|_{p_1=p}$, as defined in Davies (1977), is not continuous in the entire interval — this violates an assumption needed in Davies though our model warrants that $\{1/2\}$ is excluded.
- (ii) If $p \in (0, 1/2)$ then $\rho_{11}(p) = 0$. Here p does not appear at all and Davies' technique is not meaningful.
- (iii) If $p \in (1/2, 1)$, we face the same difficulty as in (ii). However, in (i), the test can be proposed following Davies' approach as :

$$\text{Reject } H_0 \text{ if } n^{-1/2} \left| \sum_i X_i \right| > C.$$

Distributional results do not apply since our interval for p is degenerate, i.e., $\{1/2\}$ is deleted from $(0, 1)$. Also the technique of using Lemma 3.1 of SenGupta and Chang (2000), which earlier resulted in substantial ease in proving the monotonicity of the power function of the test does no longer apply here. Nevertheless, distributional results are obtainable under both H_0 and H_1 directly. For example, cut-off points are obtained as $-\tau_{\alpha/2}$ and $\tau_{\alpha/2}$ and the power can be obtained as

$$\begin{aligned} \text{Power}(\mu, p) &= 1 - \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} P_\mu \left\{ n^{-1/2} \sum_i X_{ki} < \tau_{\alpha/2} \right\} \\ &\quad + \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} P_\mu \left\{ n^{-1/2} \sum_i X_{ki} < -\tau_{\alpha/2} \right\}, \end{aligned}$$

where $X_{k1}, X_{k2}, \dots, X_{kn}$ are i.i.d. $N(\mu, 1)$ and $X_{k,k+1}, X_{k,k+2}, \dots, X_{kn}$ are i.i.d. $N(-\mu, 1)$. This test need not possess the optimality property of

Davies' test as given in Theorem 4.2 of Davies (1977). However it is 'optimal' in the sense of L-optimality of Chapter 7 (see also SenGupta, 1991 ; Pal and SenGupta, 2000a) which clearly resembles that of Davies when the optimal test is LMP for given p_0 and is then a strong consequence.

(b) σ^2 unknown: In this case reduction by similarity (w.r.t. σ^2) can be first achieved as in II(b) and then one can pursue in the lines of III(a).

3.3. Tests for the mixing proportion

In the above situations under $H_0 : \mu = 0$, the model (3.2.1) reduced to a single normal population with mean 0. We can, however, test for 'no mixture' defining hypotheses for p , e.g., $H_0 : p = 1$ ($p = 0$), against $H : p < 1$ ($p > 0$). Under H_0 , we get a single normal population whose mean may not be specified. Suppose we want to test $H_0 : p = 1$ against $H : p < 1$. Several situations arise.

(a) $\sigma^2 = 1$: When $\mu (> 0)$ is known, simple LMP test can be found (see Durairajan and Kale, 1979) whose cut-off points can be computed at least asymptotically.

When $\mu (> 0)$ is unknown, it can be easily verified that no LMPS test exists. It is then natural to find the optimal test in the class of $C(\alpha)$ tests. Following the notations of Neyman (1959), one sees that

$$\varphi_p \equiv \left. \frac{\partial \log g}{\partial p} \right|_{p=1} = 1 - 2e^{-2\mu X} \quad \text{and} \quad \varphi_\mu \equiv \left. \frac{\partial \log g}{\partial \mu} \right|_{p=1} = X - \mu.$$

Also $\text{Cov}_{H_0}(\varphi_p, \varphi_\mu) = 2\mu$ and $\text{Var}_{H_0}(\varphi_p) = e^{4\mu^2} - 1$, so that the regression coefficient of φ_p on φ_μ under H_0 is 2μ . Again the variance of $\varphi_p - 2\mu\varphi_\mu$ under H_0 is $e^{4\mu^2} - 4\mu^2 - 1$. Observe that $\sum_{i=1}^n \varphi_\mu(X_i, \hat{\mu}) \equiv 0$, where $\hat{\mu} = \bar{X}$ is the MLE of μ under H_0 . Hence an optimal $C(\alpha)$ test rejects H_0 if

$$\frac{n^{-1/2}}{(e^{4\hat{\mu}^2} - 4\hat{\mu}^2 - 1)^{1/2}} \sum_{i=1}^n \{1 - e^{-2\hat{\mu}X_i}\} < -\tau_\alpha.$$

(b) σ^2 unknown: Whether μ is known or unknown, no reduction by similarity is possible, hence for both the cases optimal $C(\alpha)$ tests are derived. These derivations are worked out in the following

Theorem 3.3.1. (a) When μ is known and is, say μ_0 , an optimal $C(\alpha)$ test rejects $H_0 : p = 1$ against $H_1 : p < 1$ whenever

$$\frac{n^{-1/2}}{\{\exp(4\mu_0^2/\hat{\sigma}^2) - 1 - 8\mu_0^4/\hat{\sigma}^4\}^{1/2}} \sum_{i=1}^n \left\{ 1 - \exp\left(-\frac{2\mu_0 X_i}{\hat{\sigma}^2}\right) \right\} < -\tau_\alpha$$

where $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu_0)^2$.

(b) When μ is unknown, an optimal $C(\alpha)$ test rejects H_0 against H_1 whenever

$$n^{-1/2} \left(\exp(4\hat{\mu}^2/\hat{\sigma}^2) - 1 - \frac{8\hat{\mu}^4}{\hat{\sigma}^4} - \frac{4\hat{\mu}^2}{\hat{\sigma}^2} \right)^{-1/2} \sum_{i=1}^n \left\{ 1 - \exp\left(-\frac{2\hat{\mu} X_i}{\hat{\sigma}^2}\right) \right\} < -\tau_\alpha$$

where $\hat{\mu} = \bar{X}$ and $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$.

Proof. (a) It can be verified that $\varphi_p = 1 - \exp\{-2\mu_0 X/\sigma^2\}$ and $\varphi_{\sigma^2} = \frac{(X - \mu_0)^2}{2\sigma^4} - \frac{1}{2\sigma^2}$. Again by some routine algebra one sees that $\text{Var}_{H_0}(\varphi_{\sigma^2}) = \frac{1}{2\sigma^4}$; $\text{Cov}_{H_0}(\varphi_p, \varphi_{\sigma^2}) = -\frac{2\mu_0^2}{\sigma^4}$. Hence the regression coefficient of φ_p on φ_{σ^2} is $b^0 = -4\mu_0^2$. Further $\text{Var}_{H_0}(\varphi_p - b^0 \varphi_{\sigma^2}) = \exp\left\{\frac{4\mu_0^2}{\sigma^2}\right\} - 1 - \frac{8\mu_0^4}{\sigma^4}$.

Finally using the MLE $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu_0)^2$ of σ^2 under H_0 , an optimal $C(\alpha)$ test consists in rejecting H_0 whenever

$$\frac{n^{-1/2}}{\{\exp(4\mu_0^2/\hat{\sigma}^2) - 1 - 8\mu_0^4/\hat{\sigma}^4\}^{1/2}} \sum_{i=1}^n \left\{ 1 - \exp\left(-\frac{2\mu_0 X_i}{\hat{\sigma}^2}\right) \right\} < -\tau_\alpha.$$

(b) Here

$$\varphi_p = 1 - \exp\{-2\mu X/\sigma^2\}; \varphi_{\sigma^2} = \frac{(X - \mu)^2}{2\sigma^4} - \frac{1}{2\sigma^2} \text{ and } \varphi_\mu = \frac{X - \mu}{\sigma^2}.$$

The dispersion matrix of $(\varphi_p, \varphi_{\sigma^2}, \varphi_\mu)'$ under H_0 is seen to be

$$\Sigma = \begin{pmatrix} e^{4\mu^2/\sigma^2} - 1 & -2\mu^2/\sigma^4 & 2\mu/\sigma^2 \\ & 1/2\sigma^4 & 0 \\ & & 1/\sigma^2 \end{pmatrix}.$$

The partial regression coefficients are, respectively,

$$b_2^0 = -\Sigma_{12}/\Sigma_{11} = -4\mu^2 ; \quad b_3^0 = -\Sigma_{13}/\Sigma_{11} = 2\mu,$$

where Σ_{ij} is the co-factor of the (i, j) th element of Σ . It can now be verified that

$$\text{Var}_{H_0} \left\{ \varphi_P - b_2^0 \varphi_{\sigma^2} - b_3^0 \varphi_{\mu} \right\} = e^{4\mu^2/\sigma^2} - 1 - \frac{8\mu^4}{\sigma^4} - \frac{4\mu^2}{\sigma^2}.$$

Hence, as before, replacing the respective MLEs

$\hat{\mu} = \bar{X}$ and $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ of μ and σ^2 under H_0 , and noting that $\sum_{i=1}^n \varphi_{\sigma^2}(X_i, \hat{\mu}, \hat{\sigma}^2) = \sum_{i=1}^n \varphi_{\mu}(X_i, \hat{\mu}, \hat{\sigma}^2) = 0$, an optimal $C(\alpha)$ test comes out in the form: Reject H_0 if

$$n^{-1/2} \left(e^{4\hat{\mu}^2/\hat{\sigma}^2} - 1 - \frac{8\hat{\mu}^4}{\hat{\sigma}^4} - \frac{4\hat{\mu}^2}{\hat{\sigma}^2} \right)^{-1/2} \sum_{i=1}^n \left\{ 1 - e^{-\frac{2\hat{\mu}X_i}{\hat{\sigma}^2}} \right\} < -\tau_{\alpha}.$$

Hence the proof.

Chapter 4

Tests in Mixtures of Distributions in Series Forms

4.1. Introduction

This Chapter is devoted to some optimal testing problems in mixtures where one or both of the component densities are in the form of infinite series. Such is the case of a mixture obtained from the central and the corresponding noncentral distributions of, *e.g.*, χ^2 , t , F etc. distributions with noncentrality parameter, say, θ . One can also consider a mixture of bivariate Inverse Gaussian distributions (Kocherlakota, 1986), where each component density can be expressed in a series form. Another model appears in the analysis of directional data which has now-a-days been emerged as a very important branch of theoretical as well as applied statistics and is drawing attention of great many researchers. Unlike in the case of so called 'linear data', directional data need special non-trivial techniques for the analysis. Parametric inference for circular distributions is also of prime interest of many workers in this field. While circular normal (CN) distribution has extensive use as a circular distribution for many inference purposes, it has been concluded to be inadequate or "by no means conclusive because of the existence of other unimodal symmetric p.d.f.s" (SenGupta, 1998a). There are other unimodal distributions, *e.g.*, wrapped normal, wrapped cauchy etc. which are currently being widely used. These distributions are members of a large family of symmetric unimodal distributions called the family of symmetric wrapped stable (SWS) distributions (Mardia, 1972, p. 57), the densities of which are in the form of an infinite series. Circular normal distribution, although extensively used and probably the only well-known distribution to practitioners for modelling circular data, is often not appropriate for modelling real-life data. A 'suitably chosen' member of the SWS family turns out to give a better fit than a CN distribution in many such situations. Also it has been observed by Mardia (1972, p. 66)

that a CN distribution can be well approximated by a SWS distribution for some specific parameter values. Mixtures of SWS and circular uniform (CU) distributions are also very useful and important to incorporate possible outliers in the data. By taking such a mixture, the symmetry in the distribution is still retained, whereas one has more latitude in the choice of an appropriate model because of the presence of the additional parameter p , the mixing proportion.

Let $f(\theta; a, \rho, \mu_0)$ be the density function of an SWS distribution given by

$$f(\theta; a, \rho, \mu_0) = \frac{1}{2\pi} \left\{ 1 + 2 \sum_{n=1}^{\infty} \rho^n \cos n(\theta - \mu_0) \right\}. \quad (4.1.1)$$

Recall that a CN distribution with parameters κ and μ_0 has the *p.d.f.*

$$f^*(\theta; \kappa, \mu_0) = \frac{1}{2\pi I_0(\kappa)} \exp \{ \kappa \cos(\theta - \mu_0) \}$$

where $\kappa \geq 0$ and $I_0(\cdot)$ is the modified Bessel function of order 0 with purely imaginary argument. Denote by $g(\theta; a, \rho, \mu_0, p)$ the density function of a p -mixture of SWS and CU distributions, with the parameters having their usual meanings. It is this mixture family that we consider in details for our testing problems because of interesting and elegant mathematical analysis as well as practical applications, although our method may be pursued for the above class of central- noncentral mixture distributions or the bivariate Inverse Gaussian mixtures as well.

The density g , when the contaminating distribution is CU, occurs naturally in connection with experiments on perception of a group of subjects, (*e.g.*, insects) for movements towards a given direction. A goodness-of-fit test, based on Watson's U^2 -statistic incorporating a grouping correction introduced recently by Brown (1994), on Jander's ant data (Batschelet, 1981, p. 49) shows (SenGupta, 1998a) that a suitable SWS distribution gives a better fit than a CN distribution. It was also observed that a SWS-CU mixture gave a still better fit than a SWS distribution. Contamination of CU by SWS arises in human perception tests, *e.g.*, in traffic engineering, where it is generally observed that most of the individuals tend to move randomly, save a few who have rather strong perception, after undergoing some 'brain-washing' treatment. In this context a popular Indian game called 'Breaking the Pitcher' is worth-mentioning, where the player is first

shown the position of the target, then blind-folded and rotated randomly on the initial position and then asked to choose his/her own direction to break the pitcher with a stick.

The problem of testing isotropy is quite important and has received considerable attention. Several tests for this purpose under different setups (types of alternatives) exist in the literature. Beran (1968, 1969) considered this problem under a quite general set-up and derived a general form of the LMPI test. He noted that Ajne's test, Watson's U^2 test, Rayleigh's R^2 test etc. are different particular cases of the general test corresponding to different choices of the alternative hypotheses. Later Gine (1975) developed the theory of Sobolev tests as a large class of tests containing many of the known ones. Chang (1991) and SenGupta and Chang (2000) have considered locally most powerful (location) invariant (LMPI) test (with unknown μ_0) for isotropy, in terms of the hypothesis involving ρ , extensively under the model f and the LMP test under g when both p and μ_0 are known.

In this Chapter we restrict to the density g and give a new derivation of the LMPI test for the null hypothesis $H_0 : p = 0$ against $H_1 : p > 0$ assuming ρ and a to be known. This derivation, based on an expansion of the most powerful invariant test statistic in powers of p (see, e.g., Bhattacharyya and Johnson, 1969), is given in Section 4.2. It may be instructive to note the approach of Beran (1968) in its general formulation through group spaces for any 'arbitrary' density together with the explicit derivation of ours for the specific case of the general family of SWS-CU mixtures. Asymptotic distributions of the test statistic, both under the null and the alternative hypotheses, are also presented. Section 4.3 provides the global monotonicity of the power function and consistency of the test. Exact cut-off points and power values for some selected parameter combinations, obtained through extensive simulations, are presented in Chapter 11. The power computation reveals quite encouraging performances for reasonable parameter combinations even with moderate sample sizes. For the general situation when ρ is also unknown, the problem is tackled through the newly developed P^3 approach and the solution is deferred till Chapter 7. The 'non-regular' situation when ρ is known and a is unknown has been treated in Section 4.5. Some examples and summary results, in support of our proposed model, are presented in Section 4.6. These examples are also used

to demonstrate the applications of various optimal tests obtained in earlier Sections, including the Rayleigh's test. A large part of the computations of this Section has been done using the statistical package DDSTAP (Sen-Gupta, 1998b). Finally, some diagrams, also obtained by using DDSTAP, are displayed in Chapter 11. Part of this Chapter is based on SenGupta and Pal (2001).

4.2. LMPI test in the SWS-CU mixture family

The model under consideration is

$$g(\theta; a, \rho, \mu_0, p) = pf(\theta; a, \rho, \mu_0) + q(2\pi)^{-1} \quad (4.2.1)$$

where $f(\theta; a, \rho, \mu_0)$ is as given in (4.1.1). Here $0 \leq \theta < 2\pi$, $0 \leq \mu_0 < 2\pi$, $0 \leq \rho \leq 1$, $0 < a \leq 2$, $0 \leq p \leq 1$, $q = 1 - p$; ρ and a are known while μ_0 and p are unknown; μ_0 being the location parameter. It may be remarked in this context that although ρ and a are assumed to be known, in practice either or both of them may be unknown. These aspects will be discussed in subsequent Sections. Suppose $(\theta_1, \theta_2, \dots, \theta_m)$ is a random sample of size $m (\geq 2)$ from a population with density given by (4.2.1). We want to test the hypothesis $H_0 : p = 0$ against $H_1 : p > 0$. Denote by R_n^2 the quantity $\left(\sum_{j=1}^m \cos n\theta_j \right)^2 + \left(\sum_{j=1}^m \sin n\theta_j \right)^2$. The form of the LMPI test is derived in the following

Theorem 4.2.1. For known a and ρ , the LMPI test for H_0 against H_1 is given by the critical region $\omega : T = \sum_{n=1}^{\infty} (\rho^2)^{n^a} R_n^2 > C$, where the constant C is to be determined from the size condition.

Proof. First note that the problem of testing H_0 against H_1 remains invariant under the change of location $\theta_i \rightarrow \theta_i + c \pmod{2\pi}$. A set of maximal invariant statistics is $(\theta_1 - \theta_m, \theta_2 - \theta_m, \dots, \theta_{m-1} - \theta_m)$. Based on this maximal invariant the most powerful invariant test for H_0 against a fixed $p > 0$ is given by the test statistic

$$T^*(p) = \int_0^{2\pi} \prod_{i=1}^m \left[p(2\pi)^{-1} \left\{ 1 + 2 \sum_{n=1}^{\infty} \rho^{n^a} \cos n(x + \theta_i) \right\} + q(2\pi)^{-1} \right] dx. \quad (4.2.2)$$

To get the LMPI test we expand $T^*(p)$ in powers of p and consider the lowest order random term (see, e.g., Bhattacharyya and Johnson, 1969).

Now the right side of (4.2.2) is

$$\left(\frac{1}{\pi}\right)^m \int_0^{2\pi} \prod_{i=1}^m \left\{ \frac{1}{2} + p Y_i(x) \right\} dx \quad (4.2.3)$$

where

$$Y_i(x) = \sum_{n=1}^{\infty} \rho^{n^a} \cos n(x + \theta_i).$$

The coefficient of p in (4.2.3) is, apart from a multiplicative constant,

$$\sum_{i=1}^m \int_0^{2\pi} Y_i(x) dx = \sum_{i=1}^m \left[\int_0^{2\pi} \left\{ \sum_{n=1}^{\infty} \rho^{n^a} \cos n(x + \theta_i) \right\} dx \right].$$

The signs of summation and integration in the square bracket of the above expression are seen to be interchangeable by virtue of the extended version of Levi theorem for series of functions (Result D2 of Appendix D). It then easily follows that the required coefficient of p is zero. The coefficient of p^2 on the other hand, having again a multiplicative constant, is

$$\begin{aligned} & \sum_{i < j} \int_0^{2\pi} Y_i(x) Y_j(x) dx \\ &= \sum_{i < j} \int_0^{2\pi} \left[\sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \rho^{n^a + k^a} \cos n(x + \theta_i) \cos k(x + \theta_j) \right] dx \\ &= 2\pi \sum_{i < j} \sum_{n=1}^{\infty} (\rho^2)^{n^a} \cos n(\theta_i - \theta_j) \end{aligned}$$

by the same reasoning as above for interchangeability of summation and integration, followed by some algebraic manipulations. We can then write

$$T^*(p) = k_1 + k_2 p^2 \sum_{i < j} \sum_{n=1}^{\infty} (\rho^2)^{n^a} \cos n(\theta_i - \theta_j) + o_P(p^2)$$

where k_1 and $k_2(> 0)$ are constants. The critical region of the LMPI test is therefore

$$\omega^* : T^* = \sum_{i < j} \sum_{n=1}^{\infty} (\rho^2)^{n^a} \cos n(\theta_i - \theta_j) > C^* \quad (4.2.4)$$

where C^* is to be determined from the size condition. This test is obviously equivalent to ω as given in the statement of the theorem. Hence the proof. $\square$

Remark 4.2.1. It is known that the Cardioid distribution can be viewed as obtained from an SWS distribution by retaining only the first term of the infinite series occurring in the expression of its p.d.f. f . Analogous results do, therefore, hold for the LMPI test statistic in the cardioid-CU mixture family. We then have the following

Corollary 4.2.1. The LMPI test for $H_0 : p = 0$ against $H_1 : p > 0$ in the Cardioid-CU mixture family corresponds to the statistic T with $n = 1$ and therefore coincides with the Rayleigh's R^2 test.

The asymptotic null distribution of the test statistic may be obtained from Corollary 3.1 of Beran (1969) or Theorem 4.1 of Gine (1975). We however prove the following theorem which establishes the asymptotic distribution explicitly by directly appealing to the multivariate Central Limit Theorem (CLT).

Theorem 4.2.2. Under H_0 , the asymptotic distribution of $\frac{2}{m} T$ is the same as the distribution of $\sum_{n=1}^{\infty} (\rho^2)^{n^a} \chi_{(n)}^2$ where $\{\chi_{(n)}^2, n = 1, 2, \dots\}$ is a sequence of independent χ^2 variables each with 2 d.f.

Proof. Straightforward calculations show that under circular uniformity of θ

$$E(\sin n\theta) = E(\cos n\theta) = 0;$$

$$\text{Var}(\sin n\theta) = \text{Var}(\cos n\theta) = \frac{1}{2}, \text{ for each } n = 1, 2, \dots;$$

$$\text{Cov}(\sin n\theta, \sin k\theta) = \text{Cov}(\sin n\theta, \cos k\theta) = \text{Cov}(\cos n\theta, \cos k\theta) = 0;$$

for each $n, k = 1, 2, \dots; n \neq k$.

Consequently for any n , by multivariate CLT

$$m^{-\frac{1}{2}} \begin{pmatrix} \sum_i \cos \theta_i \\ \sum_i \sin \theta_i \\ \sum_i \cos 2\theta_i \\ \sum_i \sin 2\theta_i \\ \dots\dots\dots \\ \sum_i \cos n\theta_i \\ \sum_i \sin n\theta_i \end{pmatrix} \xrightarrow{\mathcal{L}} N_{2n}(\mathbf{0}, \Sigma) \text{ as } m \rightarrow \infty$$

where

$$\Sigma_{(2n \times 2n)} = \frac{1}{2} \begin{pmatrix} I_2 & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & I_2 & \dots & \mathbf{O} \\ \dots & \dots & \dots & \dots \\ \mathbf{O} & \mathbf{O} & \dots & I_2 \end{pmatrix}.$$

This shows that the limiting distribution of $\frac{2}{m}T = \frac{2}{m} \sum_{n=1}^{\infty} (\rho^2)^{n^a} R_n^2$ is the same as the distribution of $\sum_{n=1}^{\infty} (\rho^2)^{n^a} \chi_{(n)}^2$ where $\{\chi_{(n)}^2, n = 1, 2, \dots\}$ is a sequence of independent χ^2 variables each with 2 d.f.. Hence the theorem. $\square$

To obtain the asymptotic non-null distribution of T , one may appeal to the general result in Theorem 1 of Beran (1969). But the derivation of the distribution involves evaluations of some complicated integrals. We present here an alternative derivation which exploits multivariate CLT directly as in the null case.

First note that since T is invariant under change of location, one may take, without loss of generality, $\mu_0 = 0$. After some routine algebra it follows that under the mixture alternative with density $g(\theta; a, \rho, 0, p)$ and for any $n = 1, 2, \dots$,

$$E(\cos n\theta) = p\rho^{n^a}, \quad E(\sin n\theta) = 0,$$

$$\text{Var}(\cos n\theta) = \frac{1}{2} \left[1 + p\rho^{(2n)^a} \right] - p^2(\rho^2)^{n^a},$$

$$\text{Var}(\sin n\theta) = \frac{1}{2} \left[1 - p\rho^{(2n)^a} \right], \quad \text{Cov}(\cos n\theta, \sin n\theta) = 0.$$

It also follows that for any $n, l = 1, 2, \dots; n \neq l$,

$$\text{Cov}(\sin n\theta, \sin l\theta) = \frac{p}{2} \left[\rho^{|n-l|^a} - \rho^{(n+l)^a} \right],$$

$$\text{Cov}(\sin n\theta, \cos l\theta) = 0,$$

$$\text{Cov}(\cos n\theta, \cos l\theta) = \frac{p}{2} \left[\rho^{|n-l|^a} + \rho^{(n+l)^a} \right] - p^2 \rho^{n^a+l^a}.$$

For $r, s = 1, 2, \dots$, writing

$$\sigma_{rs}^{(CC)} = \text{Cov}(\rho^{r^a} \cos r\theta, \rho^{s^a} \cos s\theta),$$

$$\sigma_{rs}^{(CS)} = \text{Cov}(\rho^{r^a} \cos r\theta, \rho^{s^a} \sin s\theta),$$

$$\sigma_{rs}^{(SS)} = \text{Cov}(\rho^{r^a} \sin r\theta, \rho^{s^a} \sin s\theta)$$

and noting that $\sigma_{rs}^{(CS)} = 0 \forall r$ and s , one sees again by multivariate CLT that, under the mixture alternative

$$m^{-\frac{1}{2}} \begin{pmatrix} \rho \sum_i \cos \theta_i - mp\rho^2 \\ \rho \sum_i \sin \theta_i \\ \rho^{2a} \sum_i \cos 2\theta_i - mp(\rho^2)^{2a} \\ \rho^{2a} \sum_i \sin 2\theta_i \\ \dots\dots\dots \\ \rho^{n^a} \sum_i \cos n\theta_i - mp(\rho^2)^{n^a} \\ \rho^{n^a} \sum_i \sin n\theta_i \end{pmatrix} \xrightarrow{\mathcal{L}} N_{2n}(\mathbf{0}, \Sigma^*) \text{ as } m \rightarrow \infty$$

where

$$\Sigma^*_{(2n \times 2n)} = \begin{pmatrix} \sigma_{11}^{(CC)} & 0 & \sigma_{12}^{(CC)} & 0 & \dots & \sigma_{1n}^{(CC)} & 0 \\ 0 & \sigma_{11}^{(SS)} & 0 & \sigma_{12}^{(SS)} & \dots & 0 & \sigma_{1n}^{(SS)} \\ \sigma_{21}^{(CC)} & 0 & \sigma_{22}^{(CC)} & 0 & \dots & \sigma_{2n}^{(CC)} & 0 \\ 0 & \sigma_{21}^{(SS)} & 0 & \sigma_{22}^{(SS)} & \dots & 0 & \sigma_{2n}^{(SS)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \sigma_{n1}^{(CC)} & 0 & \sigma_{n2}^{(CC)} & 0 & \dots & \sigma_{nn}^{(CC)} & 0 \\ 0 & \sigma_{n1}^{(SS)} & 0 & \sigma_{n2}^{(SS)} & \dots & 0 & \sigma_{nn}^{(SS)} \end{pmatrix}.$$

Now applying [(ii), 6a.2] of Rao (1973, p. 387) (Result A1 of Appendix A) and taking $n \rightarrow \infty$, one can see that

$$m^{-\frac{3}{2}} \left(T - (pm)^2 \sum_{k=1}^{\infty} (\rho^4)^{k^a} \right) \xrightarrow{\mathcal{L}} N(0, \sigma^{*2}) \text{ as } m \rightarrow \infty \quad (4.2.5)$$

where

$$\sigma^{*2} = 4p^2 \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} (\rho^2)^{r^a+s^a} \sigma_{rs}^{(CC)}. \quad (4.2.6)$$

We then have the following

Theorem 4.2.3. For any $p > 0$ under the alternative hypothesis H_1 , the asymptotic distribution of T is given by (4.2.5) with σ^{*2} given by (4.2.6).

Remark 4.2.2. Exact distributions of the test statistic T under both the null and the alternative hypotheses are analytically intractable. That is why we take recourse to extensive simulations (in Section 4.4) to obtain exact cut-off points and power values of the LMPI test for different parameter and sample size combinations.

Remark 4.2.3. The asymptotic distribution of $\frac{2}{m}T$ under H_0 , as given by Theorem 4.2.2, is not convenient to carry out the test in practice. We, therefore, try to give an approximation for the above distribution. The approximation of Satterthwaite (1946) seems to be not appropriate for this purpose, as it yields very bad results for some parameter values. We, therefore, adopt a different approach based on characteristic functions. Note that under circular uniformity, the asymptotic characteristic function of $\frac{2}{m}T$ is $\varphi(t) = \left\{ \prod_{n=1}^{\infty} (1 - 2(\rho^2)^{n^2} it) \right\}^{-1}$, a first order approximation of which is $(1 - 2Kit)^{-1}$ where $K = \sum_{n=1}^{\infty} (\rho^2)^{n^2}$. This is clearly the characteristic function of $K.Z$ where Z has a chisquare distribution with 2 d.f. The values of the cut-off points obtained using this approximation are quite close to those obtained by simulations at least for some parameter combinations. For example, for $m = 20, \alpha = .05$ and for $\rho = .5$ and $a = 1.5$, the cut-off point is calculated as 16.21 which compares favorably with the simulated value 16.08 of Table 11.2.1.

4.3. Monotonicity of the power function and consistency of the test

It is convenient to work with the test statistic T^* given in (4.2.4). Monotonicity and consistency follow from the following

Theorem 4.3.1. For any fixed ρ and a , the test given in (4.2.4) possesses a monotone power function in $p \in [0, 1]$. Further the test is also consistent.

Proof. To prove the first part of the theorem we note that the p.d.f. of $\gamma_{ij} = \theta_i - \theta_j \pmod{2\pi}$, where θ_i and θ_j are independently distributed as

(4.2.1), can be written as

$$h(\gamma_{ij}; p) = \frac{1}{2\pi} \left\{ 1 + 2 \sum_{n=1}^{\infty} p^2 (\rho^2)^{n^a} \cos n\gamma_{ij} \right\}, 0 \leq \gamma_{ij} < 2\pi.$$

By Result B1 of Appendix B, for given c , $\exists$ a $\delta \in (0, \pi) \ni$

$$\sum_{n=1}^{\infty} (\rho^2)^{n^a} \cos n\gamma_{ij} > c \iff 0 < \gamma_{ij} < \delta \text{ or } 2\pi - \delta < \gamma_{ij} < 2\pi.$$

Hence for any $i, j (i < j)$

$$\begin{aligned} P(p) &= P \left\{ \sum_{n=1}^{\infty} (\rho^2)^{n^a} \cos n\gamma_{ij} > c \mid p \right\} \\ &= 2 \int_0^{\delta} h(\gamma_{ij}; p) d\gamma_{ij} \\ &= \frac{\delta}{\pi} + \frac{2p^2}{\pi} \int_0^{\delta} \sum_{n=1}^{\infty} (\rho^2)^{n^a} \cos n\gamma_{ij} d\gamma_{ij} \end{aligned}$$

and on differentiation w.r.t. p

$$\begin{aligned} \frac{\partial P}{\partial p} &= \frac{4p}{\pi} \int_0^{\delta} \sum_{n=1}^{\infty} (\rho^2)^{n^a} \cos n\gamma_{ij} d\gamma_{ij} \\ &= \frac{4p}{\pi} \sum_{n=1}^{\infty} b_n \sin n\delta \end{aligned}$$

with $b_n = \frac{1}{n} (\rho^2)^{n^a}$, by interchanging again the order of summation and integration.

Thus $P(p)$ is increasing globally in $p \in [0, 1]$ by Result B2 of Appendix B. Repeated applications of Result C2 of Appendix C on the variables

$$\sum_{n=1}^{\infty} (\rho^2)^{n^a} \cos n\gamma_{ij} \text{ for all } i, j (i < j)$$

ensure the monotonicity of the power function.

To prove consistency, it suffices (from Theorem 4 of Beran, 1969) to verify that the function $b(\theta; a, \rho, \mu_0, p)$ defined by

$$b(\theta; a, \rho, \mu_0, p) = \int_0^{2\pi} \left[f(\theta; a, \rho, x) - \frac{1}{2\pi} \right] g(x; a, \rho, \mu_0, p) dx,$$

corresponding to any density $g(\theta; a, \rho, \mu_0, p)$ under the alternative hypothesis, is non-zero.

Now

$$\begin{aligned} b(\theta; a, \rho, \mu_0, p) &= \frac{1}{\pi} \int_0^{2\pi} \left[\sum_{n=1}^{\infty} \rho^{n^a} \cos n(\theta - x) \right] \left[\frac{1}{2\pi} + \frac{p}{\pi} \sum_{n=1}^{\infty} \rho^{n^a} \cos n(x - \mu_0) \right] dx \\ &= \frac{p}{\pi} \sum_{n=1}^{\infty} (\rho^2)^{n^a} \cos n(\theta - \mu_0), \end{aligned}$$

after simplification followed by some algebraic manipulations. Thus

$$b(\theta; a, \rho, \mu_0, p) \neq 0,$$

proving the consistency of the test. Hence the theorem. $\square$

4.4. Simulation and computation

In this Section we present discussions on the exact cut-off points and power values of the LMPI test, obtained through simulations, for some selected parameter combinations. The cut-off points and power values are shown, respectively, in Table 11.2.1 and Table 11.2.2 of Chapter 11.

Simulations from a symmetric stable distribution have been done using the *RNSTA* subroutine of *IMSL*. An observation X from a symmetric stable distribution with 'scale factor' 1, when multiplied by $d^{1/\alpha}$ gives another, say Y , with scale factor d , having the characteristic function

$$\phi_Y(t) = \exp(-d|t|^\alpha).$$

Given ρ , one can take $d = -\ln \rho$ and wrap the resulting Y over $(0, 2\pi)$ to get an observation from $\text{SWS}(0, \rho, \alpha)$ distribution.

For each parameter and sample size combination, the cut-off points and the power values have been computed on the basis of 5000 and 1000 observations, respectively, on the statistic T calculated from samples generated from respective distributions. It may be noted that although the statistic T

appears in the form of an infinite series, it usually suffices to consider only a finite number of terms in practice. We have taken first 30 terms of the series for the computation of T . Power computations of Table 11.2.2 show encouraging performances for 'reasonable' parameter combinations even for samples of size 20. It is also worth noting that the power is increasing with ρ for each fixed m , a and p . This is expected because if a and p are fixed, the larger the value of ρ , the more the deviation of the density g from circular uniformity and this fact should be reflected by any reasonable test.

4.5. ρ known, a unknown : non-regular case

When ρ is known but a is unknown, we encounter the non-regular problem of having a only under the alternative. Motivated by Davies (1977, 1987), we enhance a technique of constructing optimal test for this case. To apply this technique, assume that $a \in [\epsilon, 2]$, where ϵ is a known small positive number.

Observe that for each a , under H_0

$$T'(a) \equiv \frac{2}{m} T(a) = \frac{2}{m} \sum_{n=1}^{\infty} (\rho^2)^{n^a} R_n^2 \xrightarrow{\mathcal{L}} \sum_{n=1}^{\infty} (\rho^2)^{n^a} \chi_{(n)}^2 \text{ as } m \rightarrow \infty. \quad (4.5.1)$$

where $\{\chi_{(n)}^2, n = 1, 2, \dots\}$ is a sequence of independent χ^2 variables each with 2 d.f. The test then consists in rejecting H_0 for large values of $\sup_{a \in [\epsilon, 2]} T'(a)$. Also observe that $T'(a)$ is monotonically decreasing in a and hence

$$\sup_{a \in [\epsilon, 2]} T'(a) = \frac{2}{m} \sum_{n=1}^{\infty} (\rho^2)^{n^{\epsilon}} R_n^2. \quad (4.5.2)$$

The significance probability of the test determined by the test statistic (4.5.2), however, cannot be directly obtained from the results given in Davies (1987), since the asymptotic distribution considered there is that of a *single* χ^2 . However the required significance probability can be calculated using the asymptotic distribution given in (4.5.1) with a replaced by ϵ . To obtain this, we recall (2.10) of Beran (1969). Then it follows that the required probability is

$$\sum_{n=1}^{\infty} a_n \exp \left[-t/2(\rho^2)^{n^{\epsilon}} \right] \quad (4.5.3)$$

where $a_n = \prod_{k \neq n} [1 - (\rho^2)^{k^\epsilon - n^\epsilon}]^{-1}$ and t is the observed value of $\sup_{a \in [\epsilon, 2]} T'(a)$ in (4.5.2).

4.6. Examples

We present here the analysis of two data sets, one on Jander's Ant data and the other on solar wind data considered by Philips et al. (1995) to demonstrate the applications of the LMPI and other optimal tests considered in previous Sections. We also recall (SenGupta, 1998a) that the CN distribution does not fit satisfactorily or is even unsuitable for these data whereas the SWS-CU mixture distribution gives a good fit.

The tests are applied to these data assuming that SWS-CU mixture model holds. Significance probability corresponding to the LMPI test has been computed using (4.5.2) with the determined value of a in place of ϵ . For the Davies-motivated test, the same formula has been used with $\epsilon = .1$, whereas the (asymptotic) significance probability for the Rayleigh's test has been calculated on the basis of a χ^2 distribution with 2 d.f. It may be remarked in this context that although the expression for the significance probability in (4.5.2) and that of a_n therein involve an infinite number of terms, for numerical computations it sufficed to retain at most 25 terms. Once rejection is obtained, the goodness-of-fit test is carried out using Watson's U^2 -statistic incorporating a grouping correction recently introduced by Brown (1994). The parameters for fitting CN are estimated by the m.l. method and for the SWS and SWS-CU cases, following SenGupta (1998a), the method of moments has been used for estimating μ_0 and ρ while p and a have been determined adaptively (using DDSTAP) to select the 'best' member of the respective family. In the following, although the parameters have been estimated from the samples, we have assumed, for the sake of simplicity, that the estimates are actually the 'true' values of the respective parameters (because the estimates are consistent and samples are of large sizes). For the purpose of comparison of the U^2 -values with the tabulated ones, we have used the appropriate figure corresponding to case 0 of Table 1 of Lockhart and Stephens (1985) since these figures are universal whatever the 'known' distribution we fit.

4.6.1. Jander's Ant data

Batschelet (1981, p. 49, Fig. 2) depicted the orientation of ants towards a black target when released in a round arena – an experiment originally conducted by Jander (1957). We adapt the data to construct a grouped frequency distribution of angles with 36 classes of equal widths; the total frequency being 146. The rose diagram of the data is shown in Figure 11.3.1 of Chapter 11. Summary results of the application of the optimal tests are presented in Table 11.2.3. It is seen that each of the three optimal tests leads to rejection of the null hypothesis of circular uniformity of the data. We therefore carry out the goodness-of-fit test. Summary results for fitting CN (μ_0, κ) , SWS (μ_0, ρ, a) and SWS-CU mixture (denoted by MIX (μ_0, ρ, a, p)) are presented in Table 11.2.4.

4.6.2. solar wind data

Philips et al. (1995, Fig. 5) consider the distribution of the pitch angle for suprathermal electrons at high southerly latitudes. We take the figures corresponding to -15° latitude and adapt the data to form a frequency distribution of the pitch angles with 12 classes of equal widths. The frequencies have been taken in the original scaled units. Figure 11.3.2 of Chapter 11 shows the rose diagram of the data. Although the diagram looks slightly bipolar, we present it to demonstrate an example of the situation where circular uniformity is present in the mixture in a relatively large proportion. We present the summary results in Table 11.2.5. Since here also each of the three optimal tests leads to rejection of the hypothesis of circular uniformity, we perform the goodness-of-fit tests, the results of which are given in Table 11.2.6.

Remark 4.6.1. The findings of Tables 11.2.4 and 11.2.6 suggest that mixtures of CU and SWS distributions are quite appropriate for the two sets of data under consideration. We envisage that there are many more real-life situations where the above mixture can be used to model the data and draw useful inferences.

Chapter 5

Tests for Uniparameter Hypotheses in Symmetric Multivariate Normal Mixtures

5.1. Introduction

This Chapter is devoted to optimal testing for uniparameter hypotheses in mixtures of multivariate normal distributions with component densities having equicorrelation structures for the correlation matrix. Symmetric multivariate normal (SMN) distribution with equicorrelation pattern (Bartlett and Please, 1963; Han, 1968; Rao, 1973, p. 196; Johnson and Wichern, 1992, p. 382) is a very popular model in reliability studies. However, little seems to be known about any method of construction of optimal tests in mixture models involving these distributions. In such a mixture population, the test for no mixture (contamination) reduces to the test for zero intraclass correlation coefficient, ρ , when the mixing proportion, p is known. Situations where distributions with known p are used are quite common in practice, *e.g.*, bidirectional ($p = 1/2$) circular normal distribution (Bartels, 1984) in directional data analysis; mixture of standard SMN (SSMN) distributions (Titterington *et al.*, 1985, p. 68) and mixture of SSMN and SMN distributions (Henze and Zirkler, 1990, p. 3610) in test for multivariate normality; mixture of SSMN distributions (Kocherlakota and Kocherlakota, 1981) and mixture of SSMN and SMN distributions (Srivastava and Lee, 1984) in robustness studies of estimators and tests etc. Also "... another general area where mixtures of distributions are important" is that of reliability studies for "the overall failure distribution" of a multi-component item (Everitt, 1985, p. 560). Intraclass correlation structure is also quite popular in "... many areas of applications, particularly population genetics..., reliability studies (products from the same machine...) ...

and survey sampling...” (Koch, 1983, p. 212). Here again, in many applied areas ρ , though unknown, is usually positive, *e.g.*, in split-plot experimental designs (Koch *et al.*, 1988, p. 48), finite population regression models in sample survey (Brewer and Tam, 1990, p. 423), multivariate linear models (Bai *et al.*, 1990, p. 515), efficient combination of experiments (Verrill *et al.*, 1990) etc. The SMN mixture model with $\rho > 0$ thus applies also, when a product, in batches, is acquired from two different suppliers (machines) in a known (or approximately known) proportion according to say, possibly time and cost considerations. Further applications of this model may thus be envisaged from the above examples as particular cases, *e.g.*, tests for multivariate normality as in Titterington *et al.* (1985) when $\rho \geq 0$, etc.; or as generalizations, *e.g.*, models for robustness studies — taking possibly dependent ($\rho > 0$), rather than independent, homoscedastic variables for the second component in the expression (1) of Srivastava and Lee (1984), or SMN components with $\rho_1 = 0$, $\rho_2 \geq 0$ rather than SSMN components in expression (1.1) of Kocherlakota and Kocherlakota (1981), or taking $\rho_1 = 0$, $\delta = 0$, $\rho_2 \geq 0$ in the model of Henze and Zirkler (1990).

In this Chapter we derive results for the SMN mixture family — the work on which was motivated by an earlier paper on the SSMN mixture family by SenGupta and Pal (1991). However, for the sake of completeness, we first briefly review the concepts and results associated with the SSMN mixture family. Details of derivations, proofs etc. for this family will be available from SenGupta and Pal (1991).

In Sections 5.2 and 5.3 we consider the problem of testing that the underlying population is SSMN against the alternative that it is a p -mixture of two SSMN distributions, with only one having possibly non-zero intraclass correlation coefficient, ρ . Each of the hypotheses $H_0 : \rho = 0$, $H_0 : p = 1$ and $H_0 : \rho = 0, p = 1$ simultaneously leads to the hypothesis of no mixture. So, to avoid the problem of identifiability, we might assume either $p \neq 1$ or $\rho \neq 0$. However, with the nuisance parameter p or ρ , the problem cannot be reduced by any of the principles of sufficiency, similarity or invariance. Thus, in order to get small sample exact optimal tests, we assume that the parameter not tested is known. It is obvious that the LRT is not even available in a closed form. An LMP test can be proposed (Efron, 1975) for this non-exponential mixture model. We use Efron’s statistical curvature both as a preliminary criterion for the expected performance of

an LMP test and also for choosing between the two families, p known or ρ known. In practice *a priori* information on p will be available, possibly more often than on ρ . For example, suppose that the same item is purchased in batches from two companies. Then it will be (at least approximately) known what proportion of the total purchase is made from each company. As such, the case of p known in univariate normal mixtures has received quite some attention (Butler, 1986; Newcomb, 1886 etc.). This, of course, might also be attributed to the difficulty involved when p is unknown. We demonstrate encouraging results for statistical curvature at $\rho = 0$ with variations of p in $(0,1)$. Also the test statistic simply coincides with the one for $H_0 : \rho = 0$ in the SSMN distribution and hence the exact null and non-null distributions and the cut-off points are available from Gokhale and SenGupta (1986). Thus a tremendous theoretical and computational gain is achieved. On the other hand an LMP test for $H_0 : \rho = 1$ can cause various derivational and distributional complexities (Self and Liang, 1987) due to the parameter p lying on the boundary of the parameter space $[0, 1]$. Further, we demonstrate that there exist several values of ρ for which the statistical curvature at $p = 1$ (or equivalently $q = 0$) is not defined. Hence, given a choice, an LMP test for $H_0 : \rho = 0$ is justified in favour of that for $H_0 : p = 1$.

Exact power computations, through extensive simulations, exhibit extremely encouraging results even with a sample size as small as 10 and for a non-local alternative as far as $p = 0.9$. These computations are presented in Table 11.2.8 of Chapter 11. Asymptotic normality and consistency of the test statistic are considered in Section 5.5. The results are extremely reassuring. Recall (Ghosh and Sen, 1985) that even in the ‘strongly identifiable’ case of mixture distributions and even under the null hypothesis, the asymptotic distribution of the LRT statistic is quite complicated.

In Section 5.6 we consider the more general case of a mixture of two SMN distributions, one with intraclass correlation ρ and the other with 0, and address the problem of constructing LMPS tests for the hypothesis of ‘no contamination’ under different situations concerning the knowledge of the other parameter(s). Monotonicity of the power functions is established in Section 5.7. Asymptotic normality and consistency of the test statistics are shown in Section 5.8.

5.2. LMP test in the SSMN mixture family

Let $f(\mathbf{x}|m, \sigma^2, \rho)$ denote the k -variate SMN density (Rao, 1973, p. 196), $N_k(\mathbf{M}, \sigma^2 \Sigma_\rho)$, with mean vector $\mathbf{M} = (m, m, \dots, m)'$, $\sigma^2 > 0$, $\Sigma_\rho = ((\rho + (1 - \rho)\delta_{ij}))$, δ_{ij} being the Kronecker delta and $-(k-1)^{-1} < \rho < 1$. This density can also be regarded as a density of k exchangeable normal random variables with the same marginal parameters m and σ^2 and has been considered in SenGupta (1982, 1987). Let $g_p(\mathbf{x}|m, \sigma^2, \rho) \equiv g_p$ be the density obtained as a ' p -mixture' of $f(\mathbf{x}|m, \sigma^2, 0) \equiv f_0$ and $f(\mathbf{x}|m, \sigma^2, \rho) \equiv f_\rho$, that is

$$g_p(\mathbf{x}|m, \sigma^2, \rho) = p(2\pi\sigma^2)^{-k/2} \exp \left\{ -(\mathbf{x} - \mathbf{M})'(\mathbf{x} - \mathbf{M})/2\sigma^2 \right\} \\ + q(2\pi\sigma^2)^{-k/2} |\Sigma_\rho|^{-1/2} \exp \left\{ -(\mathbf{x} - \mathbf{M})' \Sigma_\rho^{-1} (\mathbf{x} - \mathbf{M})/2\sigma^2 \right\} \quad (5.2.1)$$

where $0 < p < 1$; $q = 1 - p$, $\mathbf{M} = m \cdot \mathbf{1}$ and

$$\Sigma_\rho = \begin{pmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \dots & \dots & \dots & \dots \\ \rho & \rho & \dots & 1 \end{pmatrix}.$$

Let $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ be a random sample of size n from (5.2.1). Assume $m = 0, \sigma^2 = 1$. Here we consider testing $H_0 : \rho = 0$ against the one sided alternative $H_1 : \rho > 0$. Note that for large k , ρ should be non-negative. Usual reversal of the inequality in the definition of the critical region yields an analogous test for the alternative $\rho < 0$. We restrict to one sided alternatives because in most practical situations the sign of ρ is known. Also, such a restriction allows us to study further properties of the *LMP test*.

The structure of the LMP test (Spjøtvoll, 1968) for the mixture population (Durairajan and Kale, 1979) is based on the inequality

$$\sum_{s=1}^n \left[\frac{\partial}{\partial \rho} \log g_p(\mathbf{x}_s|m, \sigma^2, \rho) \right] \Big|_{\rho=0} > c \quad (5.2.2)$$

where c denotes a constant. Verification of regularity conditions for our problem is straightforward though tedious.

Since p is known, (5.2.2) reduces to a similar expression with g_p being replaced by f_p , *i.e.*, the test statistic coincides with that for the SSMN distribution as given in Gokhale and SenGupta (1986). The exact null distribution and the corresponding cut-off points are also available from that paper. Thus a tremendous gain is achieved here both from the theoretical and the computational aspects.

To complete the notation, let $\mathbf{X}_s = (X_{s1}, X_{s2}, \dots, X_{sk})'$, X_{si} denoting the i -th component of the vector, $\bar{\bar{X}} = \left(\sum_{s=1}^n \sum_{i=1}^k X_{si} \right) / nk$, and $\bar{X}_s$ be defined by $\bar{X}_s = \left(\sum_{i=1}^k X_{si} \right) / k$, $s = 1, 2, \dots, n$. Let

$$W = \sum_{s=1}^n \sum_{i=1}^k (X_{si} - \bar{X}_s)^2, \quad B = k \sum_{s=1}^n (\bar{X}_s - \bar{\bar{X}})^2$$

and

$$T = \sum_{s=1}^n \sum_{i=1}^k (X_{si} - \bar{\bar{X}})^2 = B + W.$$

Under f_p , $W/(1-\rho)$ is distributed as $\chi_{n(k-1)}^2$, $B/\{1+(k-1)\rho\}$ is distributed as χ_{n-1}^2 and W and B are independent (Rao, 1973, p. 198). Then the LMP test is given by:

$$\text{Reject } H_0 \text{ if } T_1 = \sum_{s=1}^n \sum_{i \neq j=1}^k X_{si} X_{sj} > c.$$

This test is unique and is thus admissible.

5.3. Statistical curvature

We discussed in Chapter 2 the role of Efron's statistical curvature, γ_θ , in an initial evaluation of the performance of an LMP test for a one-parameter non-exponential family. Here, in testing for no mixture, from a theoretical point of view we might be interested in a choice between the LMP tests for $H_0 : \rho = 0$, p known and $H_0 : p = 1$, ρ known. The latter hypothesis involves the test for a parameter on the boundary of the parameter

space and, as is well known (Self and Liang, 1987), may lead to analytical complexities. Nevertheless, here we compare the two tests using γ_θ as a criterion.

5.3.1. ${}_1\gamma_{q=0}^2$ as a function of ρ

Let $l_q \equiv \log g_p$. Then $\dot{l}_0 = r - 1$ and $\ddot{l}_0 = -(r - 1)^2$ where $r \equiv f_\rho/f_0$. Now,

$$r = |\Sigma_\rho|^{-1/2} \exp[\{-\rho/2k(1 - \rho)(1 + (k - 1)\rho)\} \\ \times \{\sum_i (x_i - \bar{x})^2 - (k - 1)(1 - \rho)(\sum_i x_i)^2\}], \quad (5.3.1)$$

$$\beta_j \equiv E_{q=0}(r^j) = (2\pi)^{-k/2} |\Sigma_\rho|^{-j/2} \int \exp \left[(-1/2) \left\{ \frac{(j - 1)\rho + 1}{1 - \rho} \sum_i x_i^2 \right. \right. \\ \left. \left. - \frac{j\rho}{(1 - \rho)(1 + (k - 1)\rho)} (\sum_i x_i)^2 \right\} \right] dx \quad (5.3.2)$$

$$= (2\pi)^{-k/2} |\Sigma_\rho|^{-j/2} \int \exp[-\frac{1}{2} \mathbf{x}' \mathbf{B}_j \mathbf{x}] d\mathbf{x}, \quad \text{say.} \quad (5.3.3)$$

where

$$\mathbf{x}' \mathbf{B}_j \mathbf{x} = \mathbf{x}' \left[\frac{(j - 1)\rho + 1}{1 - \rho} \mathbf{I}_k - \frac{j\rho}{(1 - \rho)(1 + (k - 1)\rho)} \mathbf{1}\mathbf{1}' \right] \mathbf{x}.$$

Hence from (5.3.3), when defined and when it exists, β_j is given by

$$\beta_j = (1 - \rho)^{\{k - j(k - 1)\}/2} \{1 + (k - 1)\rho\}^{-j/2} \\ \times [\{(j - 1)\rho + 1\}^k - \{(k\rho j)((j - 1)\rho + 1)^{k-1}/(1 + (k - 1)\rho)\}]^{-1/2}. \quad (5.3.4)$$

We first compute β_j and then $\alpha_j = \sum_{i=0}^j (-1)^i \binom{j}{i} \beta_{j-i}$, $j = 1(1)4$. Finally, ${}_1\gamma_{q=0}^2 = (\alpha_4 - \alpha_2)^2/\alpha_2^2 - \alpha_3^2/\alpha_2^3$.

5.3.2. ${}_1\gamma_{\rho=0}^2$ as a function of q

Let $l_\rho \equiv \log g_\rho$. Then $\dot{l}_0 = q\dot{f}_0/f_0$ and $\ddot{l}_0 = (qf_0\ddot{f}_0 - q^2\dot{f}_0^2)/f_0^2$. Let

$$T_1 = k\bar{x}^2, \quad T_2 = \sum_{i=1}^k (x_i - \bar{x})^2$$

and

$$T_3 = \frac{1}{2}k(k-1) + \frac{1}{4}\{(k-1)T_1 - T_2\}^2 - \{T_2 + (k-1)^2T_1\}.$$

Under f_0 , T_1 and T_2 are independently distributed with $T_1 \sim \chi_1^2$, $T_2 \sim \chi_{k-1}^2$. After some tedious calculations and using the fact that if $Y \sim \chi_n^2$ then

$$E(Y^r) = n(n+2)(n+4)\dots(n+2r-2),$$

we have, after some lengthy algebra,

$$E(\dot{l}_0^2) = \frac{1}{4}q^2 E\{(k-1)T_1 - T_2\}^2 = \frac{1}{2}k(k-1)q^2,$$

$$E(\dot{l}_0\ddot{l}_0) = -q^3k(k-1)(k-2)$$

and

$$\begin{aligned} E(\ddot{l}_0^2) &= q^2\left\{\frac{3}{2}k^4 - 5k^3 + \frac{13}{2}k^2 - 3k\right\} \\ &\quad + \frac{1}{16}q^4(60k^4 - 216k^3 + 300k^2 - 144k) \\ &\quad - \frac{1}{2}q^3(6k^4 - 20k^3 + 26k^2 - 12k). \end{aligned}$$

Finally

$${}_1\gamma_{\rho=0}^2 = \{2/k(k-1)\}\{(3k^2 - 7k + 6)(q^{-1} - 1)^2 + 4(k-1)\}.$$

5.3.3. Comparison and comments

Let $\gamma^2(1) \equiv {}_1\gamma_{\rho=0}^2$ and $\gamma^2(2) \equiv {}_1\gamma_{q=0}^2$. Consider $\gamma^2(2)$. Rewrite (5.3.2) in terms of $\sum_{i=1}^k (x_i - \bar{x})^2$ and $(\sum_{i=1}^k x_i)^2$ as in (5.3.1). Then the coefficient of $(\sum_{i=1}^k x_i)^2$ becomes negative for values of ρ, k and j such that

$$(k-1)(j-1)\rho^2 + (k+j-kj-2)\rho + 1 < 0, \quad (5.3.5)$$

for example, with $\rho = 0.05$, $k = 8$ and $j = 4$. So, β_j may not even be defined for such values of ρ, k and j which is seen from (5.3.4) also. Thus $\gamma^2(2)$ is not even defined for all ρ and k and hence Efron's criterion for the suitability of the LMP test for $H_0 : p = 0$ fails here.

Consider $\gamma^2(1)$. It exists for all k and q . Note that $\partial\gamma^2(1)/\partial q < 0$ for all k , so that $\gamma^2(1) \downarrow q$ and $\gamma^2(1) \rightarrow 8/k = \gamma^2(1)|_{q=1}$ as $q \rightarrow 1$. We recall (SenGupta, 1987) that $8/k$ is the statistical curvature of the curved exponential family f_ρ at $\rho = 0$. One can show that, for all k , $\gamma^2(1) \downarrow k$ also, but with $q \geq 0.5$. $\gamma^2(1) \rightarrow \infty$ as $q \rightarrow 0$. Thus the LMP test for $\rho = 0$ can be expected to perform well for all k and for all q not too small. The conservative behaviour for small q is to be naturally expected. The values of $\gamma^2(1)$ exhibited in Table 11.2.7 of Chapter 11 are quite encouraging. For example, even for $p = 0.5$ with $k = 5$, by Efron's rule one needs a 'moderate' sample size of $n \geq 50$ in order to expect that the LMP test performs reasonably well. In fact, in Section 5.4, we demonstrate that the test has extremely good power, for q not too small, with a much smaller sample size, e.g., even as small as a sample size of $n = 10$ only! We thus choose to proceed with the LMP test for $H_0 : \rho = 0$.

5.4. Exact power computations

From SenGupta and Pal (1991), the expression for exact power of the LMP test at any $\rho > 0$ is

$$P(T_1 > c | \rho) = 1 - \sum_{s=0}^n \binom{n}{s} p^s (1-p)^{n-s} P[(k-1)\chi_s^2 - \chi_{s(k-1)}^2 + (k-1)\{1 + (k-1)\rho\}\chi_{n-s}^2 - (1-\rho)\chi_{(n-s)(k-1)}^2 \leq c]. \quad (5.4.1)$$

We obtain the numerical values for the powers through simulations. The cut-off points are given by $c = c'_1 nk(k-1)$ where the exact values of c'_1 have been obtained through numerical integration and are tabulated in Gokhale and SenGupta (1986). For each n , p and ρ we generate 1000 T_1 values by generating 1000 values for each of the χ^2 variables involved. That is, for given n and k , 1000 values of each of $\chi_s^2, \chi_{s(k-1)}^2, \chi_{(n-s)}^2$ and $\chi_{(n-s)(k-1)}^2$, $s = 0, 1, \dots, n$, were generated with obvious modifications for $s = 0$ and $s = n$. Then the probability inside the summation sign on

the right hand side of (5.4.1), for each s , is obtained as the corresponding empirical c.d.f. From the first set of entries of Table 11.2.8 we note that the performance of the test is quite good, for small p , even for as small a sample size as 10. The power increases rapidly with ρ . The monotonicity of the power function has also been established analytically in SenGupta and Pal (1991). The rapid increase in power, for small p , with increase in n is noted from the second set of entries of Table 11.2.8. For large p , the power is not high as is to be expected.

5.5. Asymptotic distribution of the test statistic

Let $\tilde{\rho} = T_1/nk(k-1)$. We show below that $\tilde{\rho}$ is asymptotically normally distributed under both the null and the alternative hypotheses. Consistency will follow from this result. This result is also quite reassuring and important. Recall (Ghosh and Sen, 1985) that $-2\log\lambda$ has a complicated distribution for the 'strongly identifiable' case even under the null hypothesis. In the general case, the situation may be far worse as demonstrated by Hartigan (1985) through the simple example of $pN(0,1) + qN(\theta,1)$ where $-2\log\lambda \xrightarrow{P} \infty$. We are, of course, considering here a special case in that p ($\neq 1$) is assumed to be known. But, on the other hand, the component densities need not be of different functional forms, as is usually required for the 'strongly identifiable' case. In both these cases the hypothesis of no mixture is translated to a unique hypothesis on the parameter vector, but through different types of assumptions. Now

$$\tilde{\rho} = \sum_{s=1}^n U_s/n, \quad U_s = \sum_{i \neq j=1}^k X_{si}X_{sj}/k(k-1) = \alpha_1\chi_{\nu_1}^2 - \alpha_2\chi_{\nu_2}^2$$

where

$$\alpha_1 = \{[1 + (k-1)\rho]/k\}, \quad \alpha_2 = [(1-\rho)/k(k-1)], \quad \nu_1 = 1, \nu_2 = k-1$$

and $\chi_{\nu_1}^2$ and $\chi_{\nu_2}^2$ are independently distributed under f_ρ . Then

$$E(\tilde{\rho}) = p.0 + q.\rho = q\rho \quad \text{and} \quad \text{Var}(\tilde{\rho}) = [EU^2 - (q\rho)^2]/n.$$

Under f_ρ , we have

$$E_\rho U^2 = 3\alpha_1^2 + (k^2 - 1)\alpha_2^2 - 2(k-1)\alpha_1\alpha_2$$

$$\begin{aligned}
&= [(3k^2 - 7k + 6)k\rho^2 + 4k(k - 2)\rho + 2k]/k^2(k - 1) \\
&\equiv B(\rho)/k^2(k - 1), \text{ say.}
\end{aligned}$$

Then, under g_p , one has

$$\begin{aligned}
\text{Var}(U) &= [pB(0) + qB(\rho)]/k^2(k - 1) - (q\rho)^2 \\
&= [2kp + B(\rho)q]/k^2(k - 1) - (q\rho)^2 \equiv \sigma^2(\rho), \text{ say.}
\end{aligned} \tag{5.5.1}$$

Hence by the Central Limit Theorem, we have

Theorem 5.5.1. Under g_p , $\sqrt{n}(\tilde{\rho} - q\rho)$ is asymptotically distributed as a normal variable with mean 0 and variance $\sigma^2(\rho)$ as given in (5.5.1). Further this test is consistent.

5.6. SMN mixture family : the LMPS tests

Consider a general setup in which one wishes to test $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0, \theta \in \Omega \subset R^1$, in the presence of a nuisance parameter $\eta \in \mathcal{N} \subset R^k, k \geq 1$. Let D_α be the class of all similar level α tests of H_0 against H_1 and let $\beta_\phi(\theta, \eta)$ be the power function of a test $\phi \in D_\alpha$. Spjøtvoll (1968) presented the form of the LMPS test in some generality. For the existence of the test, it was assumed that $\partial\beta_\phi(\theta, \eta)/\partial\theta$ exists for all $\theta \in \Omega$ and $\eta \in \mathcal{N}$ for every $\phi \in D_\alpha$. However, it was observed later (Durairajan and Kale, 1982) that this condition alone does not suffice. The dominance of the test over a local interval with its one appropriate endpoint at θ_0 may be destroyed, since such an interval in general will depend on the nuisance parameter η . A second condition that for every $\phi \in D_\alpha$, the family $\{\partial\beta_\phi(\theta, \eta)/\partial\theta : \eta \in \mathcal{N}\}$ is equicontinuous at $\theta = \theta_0$ suffices (see Appendix E for the definition of equicontinuity of a family of functions). This is satisfied, *e.g.*, if $\partial^2\beta_\phi(\theta, \eta)/\partial\theta^2 \leq M < \infty, \theta \in \Omega$. For further details on equicontinuity see, *e.g.*, Dugundji (1975, p. 266). For the construction of the test the regularity conditions assumed were that the underlying distribution admits a probability density function with support independent of the parameters, the partial derivative in $\partial\beta_\phi(\theta, \eta)/\partial\theta_0$ can be passed inside the integral arising for the power function and the family of p.d.f.s possesses a boundedly complete sufficient statistic under H_0 .

The structure of the LMPS test for $H_0 : \rho = 0$ against $H_1 : \rho > 0$ in the SMN mixture population is based on the inequality

$$\sum_{s=1}^n \left[\frac{\partial}{\partial \rho} \log g_p(\mathbf{x}_s | m, \sigma^2, \rho) \right]_{\rho=0} > c(t), \quad (5.6.1)$$

where $c(t)$ generically denotes a constant depending on a fixed value t of the sufficient statistic under H_0 . Verification of regularity conditions for our problem of testing in the SMN mixture family (5.2.1) is straightforward though tedious. It needs to be pointed out for each of the three cases considered below that the LMPS test statistic is invariant with respect to the corresponding nuisance (location-scale) parameter(s) involved under the respective group of affine transformations $X \rightarrow Y$, $Y = aX + b1$, $a > 0, b \in R^1$. Hence $\beta_\phi(\theta, \eta) = \beta_\phi(\theta)$ is free of η . Thus the condition of equicontinuity becomes superfluous — the usual condition (Ferguson, 1967, p. 235) of mere continuity of $\partial \beta_\phi(\theta, \eta) / \partial \theta$ at $\theta = \theta_0$ suffices.

Since p is known, (5.6.1) reduces to a similar expression with g_p being replaced by f_p , i.e., the test statistics coincide with those for the SMN distribution. For ease of reference, the test statistics are quoted below from *Gokhale and SenGupta (1986)* wherefrom their exact null distributions and the corresponding cut-off points are also available. Thus here also a tremendous gain is achieved both from the theoretical and the computational aspects.

Recall the notations of Section 5.2. Under these notations and for any $m \in \mathcal{R}$ and $\sigma > 0$, if f_p is the true density then $W/(1-\rho)\sigma^2$ is distributed as $\chi_{n(k-1)}^2$, $B/\{1+(k-1)\rho\}\sigma^2$ is distributed as χ_{n-1}^2 and W and B are independent (Rao, 1973, p. 196). For the cases when the parameters m and σ^2 are known, without loss of generality we can assume $m = 0$ and $\sigma^2 = 1$. In each of the four cases, the critical region reduces to the form: Reject H_0 if $T_i > c$, so that it suffices to present the test statistics, T_i only.

$$\text{Case 1. } m \text{ and } \sigma^2 \text{ both known (Section 5.2): } T_1 = \sum_{s=1}^n \sum_{i \neq j=1}^k X_{si} X_{sj} \quad (5.6.2)$$

$$\text{Case 2. } m \text{ known and } \sigma^2 \text{ unknown: } T_2 = \sum_s \left(\sum_i X_{si} \right)^2 / \sum_s \sum_i X_{si}^2 \quad (5.6.3)$$

$$\text{Case 3. } m \text{ and } \sigma^2 \text{ both unknown: } T_3 = B/T \quad (5.6.4)$$

$$\text{Case 4. } m \text{ unknown, } \sigma^2 \text{ known: } T_4 = (k-1)B - W. \quad (5.6.5)$$

Remark 5.6.1. Each of the above three tests corresponding to Cases 2 through 4 is invariant with respect to the appropriate group of transformations. Since the class of invariant tests form a subclass of all similar tests, these tests are also LMPI. Recently Durairajan and Raman (1999) considered the construction of LMPI tests for the detection of compound departures from the general compound distributions. They noted that the LMPI test derived by them for the compound SMN distribution coincides with the LMPS test of our Case 3.

5.7. Monotonicity of the power functions

Monotonicity of the power functions in the above three Cases 2 through 4 are proved in the following

Theorem 5.7.1. Each of the LMPS tests for the Cases 2-4 above possesses monotone power function and is therefore unbiased.

Proof. We take up the three cases one by one.

Case 2. m known, σ^2 unknown: The critical region can be written as

$$T'_2 = \sum_{s=1}^n Y'_s > 0$$

where

$$Y'_s = \left(\sum_i X_{si} \right)^2 - c \sum_i X_{si}^2,$$

$Y'_s, s = 1, 2, \dots, n$, are i.i.d. and each Y'_s has the distribution given below :

$$\sigma^{-2} Y'_s \stackrel{d}{=} \begin{cases} k[1 + (k-1)\rho]\chi_1^2 \\ -c[(1-\rho)\chi_{k-1}^2 + \{1 + (k-1)\rho\}\chi_1^2] & \text{under } f_\rho \\ k\chi_1^2 - c[\chi_{k-1}^2 + \chi_1^2] & \text{under } f_0, \end{cases} \quad (5.7.1)$$

and

$$\begin{aligned} P(T_2 > c | \rho) &= 1 - P(T'_2 \leq 0 | \rho) \\ &= 1 - \sum_{s=0}^n \binom{n}{s} p^s (1-p)^{n-s} [G_0^{s*} * G_\rho^{(n-s)*}(0)] \end{aligned} \quad (5.7.2)$$

where $G_0(x)$ and $G_\rho(x)$ are respectively the c.d.f.s of Y'_s with distributions as specified in (5.7.1) and '*' denotes convolution. The expression in the RHS of (5.7.2) follows directly since T'_2 is a symmetric function of Y'_s 's.

Note that the last expression in the RHS of (5.7.1) is an increasing function of ρ for $\rho > 0$. This implies that for $0 \leq \rho_1 < \rho_2 \leq 1$, Y'_s is (stochastically) larger under f_{ρ_2} than under f_{ρ_1} . Hence $G_{\rho_2}(x) < G_{\rho_1}(x)$ for all x belonging to the continuity points of Y'_s . Consequently for any positive integer r , $G_{\rho_2}^{r*}(x) < G_{\rho_1}^{r*}(x)$. From (5.7.2) we get

$$\begin{aligned} P(T_2 > c | \rho_2) &= 1 - \sum_{s=0}^n \binom{n}{s} p^s (1-p)^{n-s} [G_0^{s*} * G_{\rho_2}^{(n-s)*}(0)] \\ &> 1 - \sum_{s=0}^n \binom{n}{s} p^s (1-p)^{n-s} [G_0^{s*} * G_{\rho_1}^{(n-s)*}(0)] \\ &= 1 - P(T'_2 \leq 0 | \rho_1) = P(T_2 > c | \rho_1). \end{aligned}$$

The power function is thus monotonically increasing and hence the test is unbiased.

Case 3. Both m, σ^2 unknown : The critical region can be equivalently written as

$$T'_3 = k(1-c) \sum_{s=1}^n (\bar{X}_s - \bar{\bar{X}})^2 - c \sum_{s=1}^n \sum_{i=1}^k (X_{si} - \bar{X}_s)^2 > 0. \quad (5.7.3)$$

Let us first regroup (Behboodan, 1972) the n observation vectors $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ into two groups: one of independent vectors $\mathbf{X}_{u1}, \mathbf{X}_{u2}, \dots, \mathbf{X}_{uu}$ with common density f_0 and the other of independent vectors $\mathbf{X}_{uu+1}, \mathbf{X}_{uu+2}, \dots, \mathbf{X}_{un}$ with common density f_ρ . For $u = 1, 2, \dots, n-1$ let

$$T_{u,n} = k(1-c) \sum_s (\bar{X}_{us} - \bar{\bar{X}}_u)^2 - c \sum_s \sum_i (X_{usi} - \bar{X}_{us})^2 \quad \text{where}$$

$$\mathbf{X}_{us} = (X_{us1}, X_{us2}, \dots, X_{usk})', \quad \bar{X}_{us} = \sum_{i=1}^k X_{usi}/k \quad \text{and} \quad \bar{\bar{X}}_u = \sum_{s=1}^n \bar{X}_{us}/n.$$

Let

$$\bar{\bar{X}}_{u1} = \sum_{s=1}^u \bar{X}_{us}/u, \quad T_{u1}^2 = \sum_{s=1}^u (\bar{X}_{us} - \bar{\bar{X}}_{u1})^2/u,$$

$$\begin{aligned}\bar{\bar{X}}_{u2} &= \sum_{s=u+1}^n \bar{X}_{us}/(n-u), & T_{u2}^2 &= \sum_{s=u+1}^n (\bar{X}_{us} - \bar{\bar{X}}_{u2})^2/(n-u), \\ Z_{1u}^2 &= \sum_{s=1}^u \sum_{i=1}^k (X_{usi} - \bar{X}_{us})^2 & \text{and} & & Z_{2u}^2 &= \sum_{s=u+1}^n \sum_{i=1}^k (X_{usi} - \bar{X}_{us})^2.\end{aligned}$$

Then

$$\begin{aligned}T_{u,n} &= ku(1-c)T_{u1}^2 + k(n-u)(1-c)T_{u2}^2 \\ &\quad + (ku(n-u)(1-c)/n)(\bar{\bar{X}}_{u1} - \bar{\bar{X}}_{u2})^2 - c(Z_{1u}^2 + Z_{2u}^2).\end{aligned}$$

Under f_ρ , $uT_{u1}^2 \sim (\sigma^2/k)\chi_{u-1}^2$, $(n-u)T_{u2}^2 \sim [1 + (k-1)\rho](\sigma^2/k)\chi_{n-u-1}^2$,

$$\begin{aligned}(\bar{\bar{X}}_{u1} - \bar{\bar{X}}_{u2})^2 &\sim [(1/u) + \{(1 + (k-1)\rho)/(n-u)\}](\sigma^2/k)\chi_1^2, \\ Z_{1u}^2 &\sim \sigma^2\chi_{u(k-1)}^2 & \text{and} & & Z_{2u}^2 &\sim (1-\rho)\sigma^2\chi_{(n-u)(k-1)}^2.\end{aligned}$$

So,

$$\begin{aligned}\sigma^{-2}T_{u,n} &\stackrel{d}{=} (1-c)\chi_{u-1}^2 + (1-c)\{1 + (k-1)\rho\}\chi_{n-u-1}^2 \\ &\quad + \{ku(n-u)(1-c)/n\} \\ &\quad \cdot [(1/u) + \{(1 + (k-1)\rho)/(n-u)\}](1/k)\chi_1^2 \\ &\quad - c[\chi_{u(k-1)}^2 + (1-\rho)\chi_{(n-u)(k-1)}^2].\end{aligned}\tag{5.7.4}$$

The distribution of $T_{u,n}$ under f_0 is the same as in (5.7.4) with $\rho = 0$. Moreover, under both f_0 and f_ρ , all the χ^2 variables occurring in (5.7.4) are mutually independent. Obvious modifications are to be made for $u = 0$ and n . We also note that the expression in the RHS of (5.7.4) is an increasing function of ρ for $\rho > 0$. Hence, for any two values ρ_1 and ρ_2 of ρ ($0 \leq \rho_1 < \rho_2$) we have

$$\begin{aligned}P(T'_3 > 0 \mid \rho_2) &= 1 - P(T'_3 \leq 0 \mid \rho_2) \\ &= 1 - \sum_{u=0}^n \binom{n}{u} p^u (1-p)^{n-u} P(T_{u,n} \leq 0 \mid \rho_2) \\ &> 1 - \sum_{u=0}^n \binom{n}{u} p^u (1-p)^{n-u} P(T_{u,n} \leq 0 \mid \rho_1) \\ &= P(T'_3 > 0 \mid \rho_1).\end{aligned}$$

This implies that the power function is monotonically increasing in ρ and hence the test is unbiased.

Case 4. m unknown, σ^2 known : In this case the critical region comes out in a similar form as in (5.7.3). Hence the monotonicity and unbiasedness follow by arguments similar to those in Case 3. This completes the proof of the theorem. $\square$

We obtain the numerical values for the powers with $n = 10$ through simulations. The cut-off points at $\alpha = .05$ are tabulated in Gokhale and SenGupta (1986). For each n, p and ρ , we generate 1000 T_i values by generating 1000 values for each of the χ^2 -variables involved, with necessary modifications for $u = 0$ and $u = n$. Power is obtained, then, through the empirical c.d.f. of T_i . From Tables 11.2.9, 11.2.10 and 11.2.11 we note that the performance of the test is quite good, for small p , even for as small a sample size as 10. The power increases rapidly with ρ . The monotonicity of the power function is clearly exhibited. It can be shown that, as in Case 1, here also the power increases rapidly, for small p , with increase in n . For large p , the power is not high. This is to be expected. For large p , the distribution g_p even under the alternative hypothesis $H_1 : \rho > 0$ approaches that under the null hypothesis $H_0 : \rho = 0$, and hence any reasonable test will have to suffer the consequences.

5.8. Asymptotic normality and consistency of the tests

All the three test statistics corresponding to the three cases considered above have normal distributions asymptotically and the tests are also consistent. These are shown in the following

Theorem 5.8.1. The test statistics T_2 through T_4 are asymptotically normally distributed both under H_0 and H_1 . Further the tests are consistent.

Proof. We accomplish the proof of the theorem by establishing the desired properties separately for each case.

Case 2. m known, σ^2 unknown: Rewrite T_2 as $T_2 = \bar{U}/\bar{V}$ where $\bar{U} = \sum_s U_s/n$, $U_s = \left(\sum_i X_{si} \right)^2$, $\bar{V} = \sum_s V_s/n$, $V_s = \sum_i X_{si}^2$. Then $Q_s = (U_s, V_s)'$, for $s = 1, 2, \dots, n$, are i.i.d. random vectors. Under H_1 , $E(U_s) = [1 + q(k -$

$1)\rho]k\sigma^2 = \xi_1$, say; $E(V_s) = k\sigma^2 = \xi_2$, say; and $\text{Disp}(\mathbf{Q}_s) = ((\sigma_{ij})) = \Sigma$, say, where σ_{ij} 's may be obtained by tedious calculations. By the multivariate CLT, it follows that $\sqrt{n}(\bar{\mathbf{Q}} - \xi) \xrightarrow{L} N_2(\mathbf{0}, \Sigma)$, where $\bar{\mathbf{Q}} = (\bar{U}, \bar{V})'$ and $\xi = (\xi_1, \xi_2)'$.

Thus under H_1 , $\sqrt{n}[T_2 - \{1 + q(k-1)\rho\}]$ is asymptotically distributed as a normal variable with mean 0 and variance η_ρ^2 , say, where $\eta_\rho^2 > 0$. Note that the function $h(\bar{U}, \bar{V}) = \bar{U}/\bar{V} \equiv T_2$ is totally differentiable. So the explicit form for η_ρ^2 , though not needed here, may be obtained by the Delta method (Result A2 of Appendix A) if one is still interested. Under H_0 the relevant quantities are obtained by substituting $\rho = 0$ in the corresponding above expressions. Let α henceforth denote the asymptotic size of the test.

$$\alpha \doteq P[(\bar{U}/\bar{V}) > c | H_0] \rightarrow 1 - \Phi[(c-1)/(\eta_0/\sqrt{n})] \Rightarrow c \doteq 1 + (\tau_\alpha \eta_0/\sqrt{n}),$$

τ_α being the upper α point of the standard normal variable. For large n the power of the test under any $\rho > 0$ can be written as

$$\begin{aligned} & 1 - \Phi \left[[1 + (\eta_0 \tau_\alpha / \sqrt{n}) - \{1 + q(k-1)\rho\}] / (\eta_\rho / \sqrt{n}) \right] \\ &= 1 - \Phi \left[\{\eta_0 \tau_\alpha - \sqrt{n}q(k-1)\rho\} / \eta_\rho \right] \rightarrow 1 \quad \text{as } n \rightarrow \infty \end{aligned}$$

since $q(k-1)\rho > 0$.

Case 3. m, σ^2 both unknown: Rewrite the critical region given by (5.6.4) as $T_3'' > c'$, where

$$\begin{aligned} T_3'' &= (B/n)/(W/n) \\ &= \left[k \sum_{s=1}^n (\bar{X}_s - \bar{\bar{X}})^2 / n \right] / \left[\sum_{s=1}^n \sum_{i=1}^k (X_{si} - \bar{X}_s)^2 / n \right] \\ &= (k \sum_{s=1}^n Y_s / n) / (\sum_{s=1}^n Z_s / n) = k\bar{Y} / \bar{Z}, \text{ say.} \end{aligned}$$

Note that, under both H_0 and H_1 , $\bar{X} \xrightarrow{P} m$ and $E(\bar{X}_s) = m$. Also under H_1 , $\text{Var}(\bar{X}_s) = [p\sigma^2 + q\{1 + (k-1)\rho\}\sigma^2]/k = \mu_\rho/k$, say. So $\bar{Y} \xrightarrow{P} E(\bar{X}_s - m)^2 = \mu_\rho/k$. Further $Z_s, s = 1, 2, \dots, n$ are i.i.d. random variables. Under H_1 , $Z_s \sim p\sigma^2\chi_{(1)}^2 + q(1-\rho)\sigma^2\chi_{(2)}^2$, where $\chi_{(1)}^2$ and $\chi_{(2)}^2$ are χ^2 variables

with d.f. $(k-1)$ each, so that $E(Z_s) = (1-q\rho)(k-1)\sigma^2 = \theta_\rho$, say, and $\text{Var}(Z_s) = \sigma^4(k-1)[(k+1)\{p+q(1-\rho)^2\} - (k-1)(1-q\rho)^2] = \beta_\rho^2$, say.

Then by CLT, $\sqrt{n}(\bar{Z} - \theta_\rho) \xrightarrow{\mathcal{L}} N(0, \beta_\rho^2)$. Let $h(\bar{Z}) = 1/\bar{Z}$. Then $\sqrt{n}(h(\bar{Z}) - 1/\theta_\rho) \xrightarrow{\mathcal{L}} N(0, \gamma_\rho^2)$, where γ_ρ^2 is obtained by the Delta method as $\gamma_\rho^2 = \beta_\rho^2/\theta_\rho^4$. Hence by Slutsky's theorem (see Result A3 of Appendix A), under any $\rho > 0$, $\sqrt{n}[T_3'' - \mu_\rho/\theta_\rho]$ is asymptotically distributed as a normal variable with mean 0 and variance $\tilde{\eta}_\rho^2 = \gamma_\rho^2\mu_\rho^2$. Under H_0 , the relevant quantities are obtained by substituting $\rho = 0$. Then

$$\alpha \doteq P[T_3'' > c' | H_0] \rightarrow 1 - \Phi[\sqrt{n}(c' - \mu_0/\theta_0)/\tilde{\eta}_0] \Rightarrow c' \doteq (\mu_0/\theta_0) + \tau_\alpha \tilde{\eta}_0/\sqrt{n}.$$

For large n , the power of the test under $\rho > 0$ can be written as

$$1 - \Phi[\{\sqrt{n}((\mu_0/\theta_0) - (\mu_\rho/\theta_\rho)) + \tau_\alpha \tilde{\eta}_0\} / \tilde{\eta}_\rho] \rightarrow 1 \quad \text{as } n \rightarrow \infty,$$

since $(\mu_0/\theta_0) - (\mu_\rho/\theta_\rho) = -qk\rho/[(k-1)(1-q\rho)] < 0$ for $\rho > 0$.

Case 4. m unknown, σ^2 known: Rewrite the critical region given by (5.6.5) as $T_4' > c/n$ where $T_4' = T_4/n = k(k-1)\bar{Y} - \bar{Z}$, $\bar{Y}$ and $\bar{Z}$ being as defined in Case 3 above. Let μ'_ρ, θ'_ρ and $\beta_\rho'^2$ denote the values of μ_ρ, θ_ρ and β_ρ^2 respectively, of Case 3 above with $\sigma^2 = 1$. Also let $\delta'_\rho = (k-1)\mu'_\rho - \theta'_\rho$. Hence by CLT and using Slutsky's theorem, under any $\rho > 0$, $\sqrt{n}[T_4' - \delta'_\rho]$ is asymptotically distributed as a normal variable with mean 0 and variance $\beta_\rho'^2$.

Then

$$\alpha \doteq P[T_4' > c/n | H_0] \rightarrow 1 - \Phi\left[\sqrt{n}\left\{\frac{c}{n} - \delta'_0\right\} / \beta'_0\right] \Rightarrow c \doteq \sqrt{n}\tau_\alpha\beta'_0 + n\delta'_0.$$

For large n , the power of the test under $\rho > 0$ can be written as

$$1 - \Phi[\{\tau_\alpha\beta'_0 - \sqrt{n}(\delta'_\rho - \delta'_0)\} / \beta'_\rho] \rightarrow 1 \quad \text{as } n \rightarrow \infty$$

since $d\delta'_\rho/d\rho = k(k-1)q > 0$. This proves the theorem. $\square$

Chapter 6

Higher Order $C(\alpha)$ Tests with Applications to Mixtures

6.1. Introduction

Neyman (1959) (henceforth Neyman) developed a class of asymptotic tests called the class $C(\alpha)$ of tests in connection with optimal testing problems in the presence of nuisance parameters when ‘exact optimum tests cannot be found’. A typical member of this class is formed under certain regularity conditions similar to Cramer (1946, p. 500), an assumption on the existence of locally root n consistent estimators of the nuisance parameters under the null hypothesis and the theory of regression. For a detailed account of the definitions, notations, terminology, theory and procedure of $C(\alpha)$ tests, we refer to Neyman. We, however, simplify some of the notations according to our convenience by making obvious modifications.

In some cases it may happen that optimum tests in this class do not exist. We illustrate this in Section 6.2 by an example on a mixture of normal distributions. It is, therefore, necessary to extend the concept of $C(\alpha)$ tests to second or higher order classes where optimal tests do exist and can be found. Section 6.3 deals with such a development. For a different example, also on normal mixture, we see that no optimal test even in the class of second order tests exists so that we need to go for the third order tests. In Section 6.4 we consider mixtures of the corresponding multivariate distributions and construct union-intersection ($\cup$ - $\cap$) tests of Roy (1953) for the hypotheses of no mixture. We see for the first example that the $\cup$ - $\cap$ test coincides with a modified version of the kurtosis test considered by Mardia (1970, 1974) whereas for the second example the test is equivalent to the skewness test, both used for testing normality. This Chapter is based on Pal (2000b).

6.2. Non-existence of an optimal $C(\alpha)$ test

Let X be a random observable (with support $\mathcal{X}$) having a distribution which depends on a parameter $\xi \in \Xi \subset \mathcal{R}$ and a vector of nuisance parameters $\theta \in \Theta \subset \mathcal{R}^s$. Suppose the logarithmic derivatives $\varphi_\xi(x, \theta)$ and $\varphi_{\theta_j}(x, \theta)$ of the density $p(x|\xi, \theta)$, for $j = 1, 2, \dots, s$, all evaluated under the null hypothesis $H_0 : \xi = \xi_0$, are Cramer functions as defined by Neyman. He showed that (Theorem 3) the critical region

$$Z_n^* = \frac{1}{\sqrt{n}} \sum_{i=1}^n f^*(x_i, \hat{\theta}_n) > \nu(\alpha),$$

where

$$f^*(x, \theta) = \frac{1}{\sigma_0(\theta)} \left\{ \varphi_\xi(x, \theta) - \sum_{j=1}^s a_j^0 \varphi_{\theta_j}(x, \theta) \right\},$$

a_j^0 's are the associated partial regression coefficients and $\hat{\theta}_n = (\hat{\theta}_{1n}, \hat{\theta}_{2n}, \dots, \hat{\theta}_{sn})$ is any root n consistent estimator of θ , is an optimal test of class $C(\alpha)$ with reference to the sequence of family of alternatives Γ_{1+} , provided $\sigma_0^2(\theta)$, the minimum variance under H_0 of the random quantity

$$\varphi_\xi(X, \theta) - \sum_{j=1}^s a_j \varphi_{\theta_j}(X, \theta)$$

is positive.

However, in some practical situations it may happen that $\sigma_0^2(\theta) = 0$. This occurs when there exists a linear relationship among $\varphi_\xi(X, \theta)$ and $\varphi_{\theta_j}(X, \theta)$ s with probability one. In that case no optimal $C(\alpha)$ test of Neyman exists. The following example illustrates such a situation:

Example 6.2.1. Let $p(x|\mu, \tau^2)$ be the density of a $\frac{1}{2}$ -mixture of $N(\mu, \tau^2)$ and $N(-\mu, \tau^2)$ distributions, so that

$$p(x|\mu, \tau^2) = \frac{1}{2\tau} \left[\phi\left(\frac{x-\mu}{\tau}\right) + \phi\left(\frac{x+\mu}{\tau}\right) \right] \quad (6.2.1)$$

where $\phi(\cdot)$ is the standard normal density with the corresponding c.d.f. $\Phi(\cdot)$; $\mu \geq 0$, $\tau > 0$ and both these parameters are unknown. Such a model is known as a *mirror-image* model or a *mirrored mixture* (Yamamoto

and Shinozaki, 2000) and is non-regular in the sense that the logarithmic derivative of $p(x|\mu, \tau^2)$ with respect to μ at $\mu = 0$ vanishes identically in x and τ^2 (SenGupta and Pal, 1993a).

Because of the above non-regularity, we make a unified reparametrization by the transformation $(\mu, \tau^2) \rightarrow (\xi, \theta)$ where $\xi = \mu^2$ and $\theta = \tau^2$ and consider the problem of testing $H_0 : \xi = 0$ against $H_1 : \xi > 0$ treating θ as the single nuisance parameter. One can then express the density p as

$$p(x|\xi, \theta) = \frac{1}{2\sqrt{2\pi\theta}} \exp\left(-\frac{x^2 + \xi}{2\theta}\right) \sum_{r=0}^{\infty} \frac{\xi^r x^{2r}}{\theta^{2r} (2r)!}.$$

It then follows that $\varphi_\xi(X, \theta) = \varphi_\theta(X, \theta) = \frac{X^2 - \theta}{2\theta^2}$. Hence $\sigma_0^2(\theta) = 0$.

6.3. Second and higher order $C(\alpha)$ tests

While generalizations of the $C(\alpha)$ approach in different directions have been attempted by many authors (see, *e.g.*, Buehler and Puri, 1966; Moran, 1970; Hall and Mathiason, 1990), no work seems to exist towards the development of higher order $C(\alpha)$ tests with applications to mixtures. We shall make such an attempt here.

Let $f(x, \theta)$ be an arbitrary Cramer function with $E_\xi f(X, \theta) = \mu(\xi, \theta)$. If $\mu(0, \theta) = 0$, we can use Taylor expansion to get

$$\mu(\xi, \theta) = \xi \mu'(0, \theta) + o(\xi)$$

(c.f. (56) of Neyman), where $\mu'(0, \theta)$ is the derivative of $\mu(\xi, \theta)$ w.r.t. ξ at $\xi = 0$ and is equal to the covariance of $f(X, \theta)$ and $\varphi_\xi(X, \theta)$, evaluated under $\xi = 0$.

In order to construct the class of second order $C(\alpha)$ tests for the null hypothesis $H_0 : \xi = 0$ against the alternative $H_1 : \xi > 0$, in the presence of $\theta (\in \Theta \subset \mathcal{R}^s)$ as the nuisance parameter, we take into account all normed Cramer functions $g(x, \theta)$ satisfying the condition of the existence of the integral

$$\int_{\mathcal{X}} g(x, \theta) p(x|\xi, \theta) dx$$

along with its thrice differentiability w.r.t. ξ at $\xi = 0$ under the sign of

integration. We then take

$$f(x, \theta) = \frac{1}{\sigma(g, \theta)} \left[g(x, \theta) - a^0 \varphi_\xi(x, \theta) - \sum_{j=1}^s b_j^0 \varphi_{\theta_j}(x, \theta) \right]$$

where a^0 and b_j^0 s are the partial regression coefficients of g on φ_ξ and φ_{θ_j} s respectively and $\sigma^2(g, \theta)$ is the variance, assumed to be positive, of

$$g(X, \theta) - a^0 \varphi_\xi(X, \theta) - \sum_{j=1}^s b_j^0 \varphi_{\theta_j}(X, \theta)$$

under H_0 . Clearly, one would then have $\mu'(0, \theta) = 0$ in addition to $\mu(0, \theta) = 0$. Taylor expansion now yields

$$\mu(\xi, \theta) = \frac{\xi^2}{2} \mu''(0, \theta) + o(\xi^2),$$

where

$$\mu''(0, \theta) = \left. \frac{\partial^2}{\partial \xi^2} \mu(\xi, \theta) \right|_{\xi=0} = \text{Cov}_{H_0} \left(f(X, \theta), \varphi_\xi^2 + \varphi_{\xi\xi} \right), \quad (6.3.1)$$

$\varphi_{\xi\xi}$ being the second logarithmic derivative of $p(X|\xi, \theta)$ w.r.t. ξ evaluated at $\xi = 0$. The class of second order $C(\alpha)$ tests is based on the statistic

$$Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n f(X_i, \hat{\theta}_n)$$

where $\hat{\theta}_n$ is a sequence of locally root n consistent estimators of θ .

Let us now define

$$f^*(x, \theta) = \frac{1}{\sigma_0^*(\theta)} \left[g^*(x, \theta) - a^{0*} \varphi_\xi(x, \theta) - \sum_{j=1}^s b_j^{0*} \varphi_{\theta_j}(x, \theta) \right]$$

where $g^*(x, \theta) = \varphi_\xi^2(x, \theta) + \varphi_{\xi\xi}(x, \theta)$, a^{0*} and b_j^{0*} s are, as before, the associated partial regression coefficients and $\sigma_0^{*2}(\theta)$ is the variance of the residual under consideration. We can now have the following

Theorem 6.3.1. If the estimates $\hat{\theta}_{jn}$ of the parameters θ_j , for $j = 1, 2, \dots, s$ are locally root n consistent (without being strongly consistent),

then the test of the hypothesis H_0 based on the sequence of critical regions $\{\omega_n^0\}$, where ω_n^0 is given by

$$\omega_n^0 = \left\{ (X_1, X_2, \dots, X_n) : Z_n^* = \frac{1}{\sqrt{n}} \sum_{i=1}^n f^*(X_i, \hat{\theta}_n) > \nu(\alpha) \right\},$$

is an optimal test of the class of second order $C(\alpha)$ tests with reference to the sequences of alternatives $\{\xi_n\}$ for which $\xi_n^2 \sqrt{n}$ remains bounded.

Proof. If $\xi_n^2 \sqrt{n}$ remains bounded as $n \rightarrow \infty$, an argument similar to Theorem 2 of Neyman establishes that the asymptotic distribution of Z_n^* is normal with mean $\mu''(0, \theta) \cdot \frac{\xi_n^2}{2} \sqrt{n}$ and variance unity. In view of (6.3.1), the theorem follows in the lines of Theorem 3 of Neyman. $\square$

Illustration. Let us now apply Theorem 6.3.1 to the example considered in Section 6.2. Note that, under the new parametrization, $\varphi_{\xi\xi} = -\frac{X^4}{6\theta^4}$.

Recall that $\varphi_\xi = \varphi_\theta = \frac{X^2 - \theta}{2\theta^2}$. Hence $\varphi_{\xi\xi} + \varphi_\xi^2 = \frac{X^4}{12\theta^4} - \frac{X^2}{2\theta^3} + \frac{1}{4\theta^2}$. Again it can be verified that $\text{Cov}_{H_0}(\varphi_{\xi\xi} + \varphi_\xi^2, \varphi_\xi) = 0$, so that the associated regression coefficient a^{0*} is zero. Consequently $b^{0*} = 0$. It therefore remains to evaluate $\sigma_0^*(\theta)$. After a little algebra it turns out that $\sigma_0^{*2}(\theta) = \frac{1}{6\theta^4}$.

Also, the maximum likelihood estimator of θ under H_0 is $\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n X_i^2$, which is also root n consistent. Replacing this estimator in Z_n^* of Theorem 6.3.1 it follows, after simplification, that an optimal test is based on the critical region

$$\omega_n^0 = \left\{ (X_1, X_2, \dots, X_n) : Z_n^* = \sqrt{6n} \left(\frac{m'_4}{12m_2'^2} - \frac{1}{4} \right) > \nu(\alpha) \right\}$$

where $m'_r = \frac{1}{n} \sum_{i=1}^n X_i^r$, the r -th raw moment about origin.

It is also seen that the asymptotic power of the above optimal test against any sequence of alternatives $\{\xi_n\}$ and any θ is given by

$$1 - \Phi \left(\nu(\alpha) - \frac{\sqrt{n}\xi_n^2}{2\sqrt{6}\theta^2} \right).$$

Remark 6.3.1. It is clear from the above discussion that the class $C(\alpha)$ of asymptotic tests, as developed by Neyman, can be termed as the class of first order $C(\alpha)$ tests. When an optimal test in the said class fails to exist, it may happen that there is one in the class of second order tests, which is essentially a subclass of the former. Optimality, however, holds with reference to a different sequence of alternatives of a higher order. Also, as the following example shows, no optimal test even among the class of second order $C(\alpha)$ tests may exist. It is then envisaged that higher order $C(\alpha)$ tests can be analogously defined under more stringent conditions and an optimal test in such a class can be constructed by an appeal to the above procedure.

Example 6.3.1. Consider now the mixture density

$$p(x|\xi, \theta) = \frac{\pi}{\theta_2^{1/2}} \phi\left(\frac{x - \theta_1}{\theta_2^{1/2}} + \xi\right) + \frac{1 - \pi}{\theta_2^{1/2}} \phi\left(\frac{x - \theta_1}{\theta_2^{1/2}}\right) \quad (6.3.2)$$

where π ($0 < \pi < 1$) is known; ξ ($-\infty < \xi < \infty$) is the unknown parameter of interest and $\theta = (\theta_1, \theta_2)$ is the vector of nuisance parameters with $-\infty < \theta_1 < \infty$ and $0 < \theta_2 < \infty$. Observe that (6.3.2) is a reparametrized version of the mixture density

$$p(x|\gamma_1, \gamma_2, \delta^2) = \frac{\pi}{\delta} \phi\left(\frac{x - \gamma_1}{\delta}\right) + \frac{1 - \pi}{\delta} \phi\left(\frac{x - \gamma_2}{\delta}\right) \quad (6.3.3)$$

with $\theta_1 = \gamma_2$, $\theta_2 = \delta^2$ and $\xi = (\gamma_2 - \gamma_1)/\delta$. Recall model (iii) of Subsection 2.4.3 of Chapter 2 in this context.

One can verify that

$$\varphi_\xi = -\pi Y; \quad \varphi_{\theta_1} = Y/\theta_2^{1/2} \text{ and } \varphi_{\theta_2} = \frac{1}{2\theta_2^{1/2}}(Y^2 - 1)$$

where $Y \equiv Y(X, \theta) = (X - \theta_1)/\theta_2^{1/2}$. Clearly the partial regression coefficient of φ_ξ on φ_{θ_1} is $-\pi\theta_2^{1/2}$, whereas that of φ_ξ on φ_{θ_2} is 0. Hence the residual variance is again 0 and no optimal test in the class of first order tests exists. As regards the optimal second order test, the key function is $\varphi_{\xi\xi} + \varphi_\xi^2 = \pi\{Y^2 - 1\}$ which is to be regressed on $Y/\theta_2^{1/2}$ and $\frac{1}{2\theta_2^{1/2}}(Y^2 - 1)$. The residual variance is again 0 and no optimal second order $C(\alpha)$ test

exists. Failure of the existence of the optimal tests of first two orders naturally leads us to consider third order tests. To achieve this end we make the obvious additional assumptions regarding the choice of the Cramer functions in this case. The key function here, analogous to g^* of the second order class, is

$$g^*(X, \theta) = \varphi_{\xi\xi\xi} + 3\varphi_{\xi\xi}\varphi_{\xi} + \varphi_{\xi}^3,$$

$\varphi_{\xi\xi\xi}$ being the third logarithmic derivative of $p(x|\xi, \theta)$ w.r.t. ξ evaluated at $\xi = 0$. After some simplification we get $g^* = \pi(3Y - Y^3)$. To construct the function f^* we have to find the multiple regression equation of g^* on φ_{ξ} and $\varphi_{\xi\xi} + \varphi_{\xi}^2$ and then take the residual. Since all the partial regression coefficients turn out to be 0 here, the required residual is g^* itself. By the same reasoning the residual of g^* after taking multiple regression on φ_{θ_1} and φ_{θ_2} is again g^* . Also $\text{Var}_{H_0}(g^*) = 6\pi^2$. Hence using the MLEs $\hat{\theta}_{1n} = \bar{X}$ and $\hat{\theta}_{2n} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ of θ_1 and θ_2 respectively under H_0 , an optimal third order $C(\alpha)$ test for H_0 against H_1 is given by the critical region

$$\omega_n^{0*} = \left\{ (X_1, X_2, \dots, X_n) : \sqrt{n/6} |g_1| > \nu(\alpha/2) \right\}$$

where g_1 is the usual measure of skewness of the data. Observe that this coincides with the so called skewness test and has been used for a long time for testing normality (see, *e.g.*, D'Agostino and Pearson, 1973; D'Agostino and Pearson, 1974; Mendell *et al.*, 1993).

Remark 6.3.2. When $\pi = 1/2$, the mixture distribution is symmetric for any value of ξ and consequently the above skewness test will fail to detect departure from normality. In this case we suggest that one should use both skewness and kurtosis tests in the lines of Mardia (1974).

6.4. Testing in multivariate models

Consider the multivariate models corresponding to (6.2.1) and (6.3.3). These are, respectively

$$p(\mathbf{x}|\mu, T) = \frac{1}{2} \{N_k(\mathbf{x}|\mu, T) + (1 - \pi)N_k(\mathbf{x}|-\mu, T)\} \quad (6.4.1)$$

and

$$p(\mathbf{x}|\gamma_1, \gamma_2, \Delta) = \pi N_k(\mathbf{x}|\gamma_1, \Delta) + (1 - \pi)N_k(\mathbf{x}|\gamma_2, \Delta) \quad (6.4.2)$$

where, in (6.4.1), μ is an unknown k component vector whereas T is an unknown positive definite symmetric matrix of order k . Similarly in (6.4.2), γ_1 and γ_2 are unknown vectors of k components and Δ is an unknown positive definite symmetric matrix. In both (6.4.1) and in (6.4.2), $N_k(\cdot|\cdot, \cdot)$ is the density function of a k -variate normal distribution with appropriate parameters. For testing the hypothesis of no mixture in each of the models, we can conveniently combine the optimal tests of the univariate problems of Sections 6.2 and 6.3 by the union-intersection ($\cup$ - $\cap$) principle of Roy (1953). A pertinent result in this context is the following lemma which establishes Cramer-Wold type theorem for mixtures of normal distributions.

Lemma 6.4.1. (Pal, 2000b) $X \sim \pi N_k(\mu_1, \Sigma_1) + (1 - \pi)N_k(\mu_2, \Sigma_2)$ iff

$$l'X \sim \pi N(l'\mu_1, l'\Sigma_1 l) + (1 - \pi)N(l'\mu_2, l'\Sigma_2 l) \text{ for every } l \neq 0.$$

Proof. Follows from standard arguments through characteristic functions. $\square$

We now construct $\cup$ - $\cap$ tests for 'no mixture' in each of the above models. We consider (6.4.1) and the treatment for the other will follow analogously. Observe that

$$H_0 : \mu = 0 \equiv \bigcap_{l \neq 0} \{H_{0l} : l'\mu = 0\}.$$

Write $Z_i = l'X_i$ for each $i = 1, 2, \dots, n$. Using Lemma 6.4.1 and noting that the distribution of $m'_4(Z)/m'^2_2(Z)$ does not depend on l under H_{0l} , it follows that the $\cup - \cap$ test rejects H_0 if

$$\sup_{l \neq 0} \sum_{i=1}^n \left(\frac{l'X_i X'_i l}{l'Al} \right)^2 > c$$

where $c = c(n)$ generically denotes a constant such that the overall size of the test is α and $A = \sum_{i=1}^n X_i X'_i$ which is positive definite with probability

1 provided $n \geq k$ (this is a consequence of Theorem 2.1 of Eaton and Perlman, 1973). But

$$\begin{aligned}
\sup_{l \neq 0} \sum_{i=1}^n \left(\frac{l' X_i X_i' l}{l' A l} \right)^2 &\leq \sum_{i=1}^n \sup_{l \neq 0} \left(\frac{l' X_i X_i' l}{l' A l} \right)^2 \\
&= \sum_{i=1}^n \left\{ \sup_{l \neq 0} \frac{l' X_i X_i' l}{l' A l} \right\}^2 = \sum_{i=1}^n (X_i' A^{-1} X_i)^2 \\
&= \sum_{i=1}^n \left\{ \inf_{l \neq 0} \frac{l' X_i X_i' l}{l' A l} \right\}^2 \leq \sum_{i=1}^n \inf_{l \neq 0} \left(\frac{l' X_i X_i' l}{l' A l} \right)^2 \\
&\leq \inf_{l \neq 0} \sum_{i=1}^n \left(\frac{l' X_i X_i' l}{l' A l} \right)^2.
\end{aligned}$$

This implies that

$$\sup_{l \neq 0} \sum_{i=1}^n \left(\frac{l' X_i X_i' l}{l' A l} \right)^2 = \sum_{i=1}^n (X_i' A^{-1} X_i)^2.$$

Hence the above test reduces to:

Reject H_0 if

$$\frac{1}{n} \sum_{i=1}^n (X_i' S^{-1} X_i)^2 > c \quad (6.4.3)$$

where $S = A/n$. It may be instructive to compare the test statistic given by (6.4.3) with the expression (3.12) for multivariate kurtosis of Mardia (1970). The cut-off point c can be found asymptotically using the limiting distribution (3.20) of Mardia (1970) with $b_{2,p}$ playing the role of the test statistic of (6.4.3) and p being replaced by k .

Proceeding in a similar fashion with the model (6.4.2) one sees that the $\cup - \cap$ test for the null hypothesis H_0 : 'no mixture' is given by :

Reject H_0 if

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \left\{ (X_i - \bar{X})' S^{*-1} (X_j - \bar{X}) \right\}^3 > c \quad (6.4.4)$$

where, as before, c satisfies the overall size condition and $S^* = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})'$. That S^* also is positive definite with probability 1 follows

from, *e.g.*, Johnson and Wichern (1992, pp. 110-111) (see also Theorem 3.4 of Eaton and Perlman, 1973). Recall that the test (6.4.4) is based on the measure of multivariate skewness $b_{1,p}$ as defined by Mardia (1974) (see also Mardia, 1970). An asymptotic expression for c can now be found from (2.26) of Mardia (1970). An algorithm for calculating $b_{1,p}$ can be found in Mardia and Zemroch (1975).

Remark 6.4.1. (i) As in the univariate case, the test (6.4.4) should be used along with the kurtosis test of Mardia (1974) when $\pi = 1/2$.

(ii) The above $\cup - \cap$ tests are derived from the corresponding univariate tests which are optimal in the sense of having the maximum asymptotic local power among tests of certain classes. No such optimality property, however, is known for these multivariate tests. Nevertheless, (6.4.3) and (6.4.4) turn out to be objective procedures for testing multivariate normality if one restricts to multivariate models (6.4.1) and (6.4.2). This may provide some insight into the relevance of using skewness or kurtosis tests as satisfactory procedures for testing multivariate normality.

Chapter 7

The P^3 , L-optimal and Estimating Function Approaches

7.1. Introduction

This Chapter deals with optimal testing problem for the hypothesis of 'no mixture' in two-component mixture distributions (with or without nuisance parameter(s)) in which the parameter of interest and the mixing proportion both are unknown. For these models, the likelihood ratio (LR) tests are very cumbersome and cannot be obtained in a closed form. Also usual asymptotic properties of this test do not hold. Hartigan (1985) showed that for the simple example of a mixture of two normal distributions, $-2 \log \lambda$, where λ is the LR criterion, converges to infinity in probability under the null hypothesis as the sample size grows indefinitely. Even in 'strongly identifiable' mixtures, Ghosh and Sen (1985) have demonstrated that the asymptotic distribution of the LR criterion is based on a certain functional of the supremum of a Gaussian process (see Chapter 1). Under a general setup in the presence of nuisance parameters, Aitkin and Rubin (1985) proposed to work with the ratio of *integrated likelihoods* taking into consideration a prior distribution of the mixing proportion(s), and thereby eliminating this (these) nuisance parameter(s). But then, as shown by Quinn *et al.* (1987), a condition needed to establish the asymptotic distribution of the test statistic does not hold.

Durairajan (1980) addressed the problem of testing the hypothesis of no mixture in two-component mixtures of distributions in some generality. When both p and the parameter (say θ) are unknown, he obtained LMPI and LMPS tests for hypotheses involving p , treating θ as the nuisance parameter. The other problem, *i.e.*, testing for θ with p as nuisance parameter can not, usually, be tackled by the Neyman-Pearson theory since p can not be eliminated by any of the principles of sufficiency, similarity or invariance.

In the following development we focus our attention to locally optimal procedures by considering what are called 'L-optimal' tests introduced by SenGupta (1991). An L-optimal test, in the context of mixture models, is one that is locally best (optimal) (LMP or LMPU or LMPI or LMPS, as the situation may be) for testing the parameter for each *given* value of the mixing proportion p . Note that SenGupta's definition of L-optimality pertains to testing for 'no contamination' in situations where the component distributions belong to a one-parameter family, with one component distribution being completely specified, so that the criterion of L-optimality would demand the test to be LMP for each given p . Likewise we may use the term 'L-optimal invariant' or 'L-optimal similar' for tests when nuisance parameters are present and the problem is reduced by invariance or similarity. The power of an L-optimal test, therefore, matches that of the corresponding locally best test for the parameter for each given p .

The idea is to first recast the original several parametric hypotheses representation of the null hypothesis in terms of only a single hypothesis involving an appropriately chosen parametric function. Such a parametric function is termed as *Pivotal Parametric Product* (P^3 in short). After an appropriate P^3 is chosen, the test is constructed on the basis of a suitable (usually unbiased and/or consistent) estimator of it. We may call such a test a P^3 test. The (locally) optimal P^3 test, if exists and is unique, can then be seen to be the L-optimal test as defined by SenGupta (1991). An L-optimal test is admissible, *i.e.*, there does not exist any other test which performs at least equally well (in terms of power function) at all the points under the alternative hypothesis and actually better than this test at some point(s). Furthermore, this test is consistent.

The present Chapter is organized in the following way. In Section 7.2 we consider the case when there is no nuisance parameter and give a set of sufficient conditions on the class of densities such that the associated class of mixtures admits an appropriate P^3 . We then characterize a general form of the appropriate P^3 and the structure of the corresponding L-optimal test. It is observed that the class of densities of the one-parameter exponential family of distributions satisfies all these conditions. Section 7.3 deals with *an integrated likelihood approach and we show that the LMP test for the parameter based on an integrated likelihood coincides with the L-optimal test*. In Section 7.4 we discuss the applications of mixtures in reliability

theory in details and note the existence of a rich class of mixtures of reliability distributions for which the forms of the P^3 s and the L-optimal tests can be conveniently derived by an appeal to the general result of Section 7.2. We also illustrate this result through several distributions useful in reliability theory — although many more distributions, from different fields of applications, can be brought under the purview of this result. For each distribution, we show that a simple characterization of P^3 is possible and the resulting L-optimal test is also very simple and elegant, both in form and in application (SenGupta and Pal, 1995). In Section 7.5 we consider the estimating function approach and establish an asymptotic optimality property of the P^3 tests among a class of tests based on linear unbiased estimating functions. In Section 7.6 we illustrate the application of our procedure through two examples — one on a real-life data set (Barnett and Lewis, 1994, p. 253) on crushed rock and the other on a simulated sample drawn from a mixture of two lognormal distributions. Finally, in Section 7.7 we deal the situation, with examples, where nuisance parameters are present in the model. Part of this Chapter is based on Pal and SenGupta (2000a, 2000b).

In what follows we shall denote by $E_{f_\theta}(\cdot)$ and $\text{Var}_{f_\theta}(\cdot)$ the expectation and variance, respectively, of a random variable taken with respect to the density $f(x|\theta)$.

7.2. The P^3 approach and L-optimal tests

Consider the mixture model with density

$$g(x|p, \theta) = pf(x|\theta) + (1 - p)f(x|\theta_0) \quad (7.2.1)$$

where $0 \leq p \leq 1$, $\theta \in \Theta$, an interval of the real line; both p, θ are unknown and θ_0 is a known point of Θ . The density $f(x|\theta)$ is assumed to be sufficiently ‘regular’. We want to test the null hypothesis H_0 : ‘no contamination’ against the alternative H_1 : ‘there is contamination’. Under the above setup, the null hypothesis of no contamination translates to the union of three parametric hypotheses : $[H_{01} : p = 0 \cup H_{02} : \theta = \theta_0 \cup H_{03} : p = 0 \text{ and } \theta = \theta_0]$.

The main idea of the P^3 approach is to characterize a single parametric function $\eta \equiv \eta(p, \theta, \theta_0)$ so that $\eta = 0$ holds iff H_0 is true. For example one

may take $\eta = p(\theta - \theta_0)$. Such a parametric function is called a *Pivotal Parametric Product* (SenGupta, 1991). Clearly several such characterizations are possible. We shall choose η so as to ensure that L-optimal test for the hypothesis $H'_0 : \eta = 0$ (which is now equivalent to H_0) can be constructed based on an unbiased and consistent estimator of η .

L-optimal tests are generally very simple in form and the cut-off points can be computed easily, at least for large samples, either analytically or by simulations.

Denote by $\mathcal{G}$ the class of density functions g , given by (7.2.1), of mixture distributions obtained by restricting f to a certain class $\mathcal{F}$ of the component density functions. The following lemma gives a set of general conditions on $\mathcal{F}$ under which each member of $\mathcal{G}$ admits a P^3 , along with its appropriate general form and the structure of the corresponding L-optimal test for members of $\mathcal{G}$. Let $X_1, \dots, X_n$ be n i.i.d. observations drawn from a population with density g .

Lemma 7.2.1. Let $\mathcal{F} = \{f\}$ be the class as referred to above with any member f of this class being a one-parameter density function, with respect to an appropriate σ -finite measure μ , of a possibly multidimensional random variable X . Assume that f satisfies the following conditions :

- (C1) The parameter θ belongs to the parameter space Θ which is a non-degenerate open, semi-open or closed interval of the real line containing θ_0 as an interior or a boundary point.
- (C2) The support $\mathcal{X}$ is independent of the parameter θ .
- (C3) The (one- or two-sided) derivative $\partial f(x|\theta)/\partial\theta|_{\theta_0}$ exists and is finite for all $x \in \mathcal{X}$.

$$(C4) \quad E_{f_{\theta}} \left(\left. \frac{\partial \log f(X|\theta)}{\partial \theta} \right|_{\theta_0} \right)^2 < \infty \text{ for all } \theta \in \Theta.$$

$$(C5) \quad E_{f_{\theta}} \left(\left. \frac{\partial \log f(X|\theta)}{\partial \theta} \right|_{\theta_0} \right) \stackrel{\text{def}}{=} \gamma(\theta, \theta_0) = 0 \text{ if and only if } \theta = \theta_0.$$

Then $\eta = p\gamma(\theta, \theta_0)$ or any monotone function of it may serve as an appropriate P^3 . A test appropriately based on an unbiased and consistent estimator T of η will be L -optimal for testing $H'_0 : \eta = 0$ against either of the one-sided alternatives, if T coincides, a.e., with the average score statistic $\frac{1}{n} \sum_{i=1}^n \left. \frac{\partial \log f(X_i|\theta)}{\partial \theta} \right|_{\theta=\theta_0}$.

Proof. By (C4) it follows that $\int \left| \frac{\partial \log f(x|\theta)}{\partial \theta_0} \right| f(x|\theta_0) d\mu < \infty$ and hence the derivative of the left side of $\int f(x|\theta) d\mu = 1$ w.r.t. θ at θ_0 can be passed inside the sign of integration. Consequently $E_{f_{\theta_0}} \left[\frac{\partial \log f(X|\theta)}{\partial \theta_0} \right] = 0$. Hence

$$\begin{aligned} E_{g_\theta} \left(\left. \frac{\partial \log f}{\partial \theta} \right|_{\theta=\theta_0} \right) &= p E_{f_\theta} \left(\left. \frac{\partial \log f}{\partial \theta} \right|_{\theta=\theta_0} \right) + (1-p) E_{f_{\theta_0}} \left(\left. \frac{\partial \log f}{\partial \theta} \right|_{\theta=\theta_0} \right) \\ &= p E_{f_\theta} \left(\left. \frac{\partial \log f}{\partial \theta} \right|_{\theta=\theta_0} \right) \\ &= p\gamma(\theta, \theta_0) \\ &= \eta. \end{aligned}$$

If T coincides, a.e., with the average score statistic, then the test appropriately based on T is LMP for $H_0 : \theta = \theta_0$ against either of the one-sided alternatives for each given $p (> 0)$, by the Neyman-Pearson lemma and hence is L -optimal for testing $H'_0 : \eta = 0$. Moreover T is unbiased for η and, by virtue of (C4), has finite variance which tends to 0 as $n \rightarrow \infty$. T is therefore consistent for η . $\square$

Corollary 7.2.1. Consider the one-parameter exponential family of distributions given by the density (w.r.t. a σ -finite measure μ) in the canonical form

$$h(x|\theta) = \exp[\theta W(x) - A(\theta)].$$

Then $h \in \mathcal{F}$ and η in this case is given by $p(A'(\theta) - A'(\theta_0))$.

Proof. It can be verified that h satisfies conditions (C1) – (C4) of Lemma 7.2.1 (see, e.g., Lehmann, 1983, p. 119). Also it is easy to check that

$\gamma(\theta, \theta_0) = A'(\theta) - A'(\theta_0)$ and since $A''(\theta) = \text{Var}_{h_\theta}(W(X)) > 0$, $A'(\theta)$ is strictly increasing in θ . Therefore (C5) is also satisfied. $\square$

Remark 7.2.1. (C5) of Lemma 7.2.1 is the condition of equivalence of the hypotheses $H_{02} : \theta = \theta_0$ and $H_0'' : \gamma(\theta, \theta_0) = 0$. This condition is needed because the equation $\gamma(\theta, \theta_0) = 0$ may have multiple solutions for θ , as is seen from the following example :

Let $f(x|\theta)$ be the density of a von Mises distribution given by

$$f(x|\theta) = [2\pi I_0(\kappa)]^{-1} \exp [\kappa \cos(x - \theta)]$$

with $0 \leq x < 2\pi$; $0 \leq \theta < 2\pi$; $\kappa (> 0)$ being known (Mardia, 1972. p. 57). Take $\theta_0 = 0$. Then

$$\left. \frac{\partial \log f}{\partial \theta} \right|_{\theta_0} = \kappa \sin x \quad \text{and} \quad E_{f_{\theta_0}} \left[\left. \frac{\partial \log f}{\partial \theta} \right|_{\theta_0} \right] = \xi(\kappa) \sin \theta \quad \text{where} \quad \xi(\kappa) \neq 0.$$

The right side, when equated to 0, yields two solutions for θ , viz., 0 and π .

7.3. Integrated likelihood approach and the optimal P^3 test

Let $L(\theta, p)$ denote the likelihood function of θ and p and let $\pi(p)$ be the density of a prior distribution of p (with respect to Lebesgue measure on $[0,1]$). The integrated likelihood of θ , $\tilde{L}(\theta)$, is then obtained by integrating $L(\theta, p)$ with respect to $\pi(p)dp$ (see Aitkin and Rubin, 1985), so that

$$\begin{aligned} \tilde{L}(\theta) &= \int_0^1 L(\theta, p) \pi(p) dp \\ &= \int_0^1 \prod_{i=1}^n \{p f(x_i|\theta) + (1-p) f(x_i|\theta_0)\} \pi(p) dp. \end{aligned}$$

$\tilde{L}(\theta)$ can now be used to construct the LMP test for a hypothesis involving θ . Note that

$$\begin{aligned} \left. \frac{d \log \tilde{L}(\theta)}{d\theta} \right|_{\theta_0} &= \left[\frac{1}{\tilde{L}(\theta)} \frac{d\tilde{L}(\theta)}{d\theta} \right]_{\theta_0} \\ &= \left[\frac{1}{\tilde{L}(\theta)} \frac{d}{d\theta} \int_0^1 \prod_{i=1}^n \{p f(x_i|\theta) + (1-p) f(x_i|\theta_0)\} \pi(p) dp \right]_{\theta_0} \quad (7.3.1) \end{aligned}$$

Assume that $\partial L(\theta, p)/\partial \theta$ is continuous on $\Theta \times [0, 1]$, so that the derivative can be passed under the integral sign in (7.3.1) (see Result D1 of Appendix D). We then have, after some simplification,

$$\left. \frac{d \log \tilde{L}(\theta)}{d\theta} \right|_{\theta_0} = K \sum_{i=1}^n \frac{f'(x_i|\theta_0)}{f(x_i|\theta_0)}$$

where

$$K = \int_0^1 p\pi(p)dp > 0.$$

Summarizing the above, we have the following

Theorem 7.3.1. The locally most powerful test for $H_0 : \theta = \theta_0$ against either of the one-sided alternatives, based on the integrated likelihood $\tilde{L}(\theta)$, is equivalent to the optimal P^3 test.

7.4. Mixtures in reliability analysis and L-optimal tests

Mixtures of distributions are recently receiving attractions in life-testing and survival analysis. Titterington *et al.* (1985, p. 20) give a detailed account of references of works on mixture models in failure-time data. In a recent paper Block and Joe (1997) considered mixtures of lifetime distributions in the study of tail behavior of failure rate functions. For mixtures of exponential distributions, they have remarked (p. 269) “ For electronic components, most of the population might be exponential, with long lives, while a small percentage often have an exponential distribution with short lives ”. Mukherjee and Chatterjee (1987) considered parameter estimation based on truncated samples from mixtures of distributions in the exponential family. Many more examples abound from which it is clear that mixtures of members of the exponential family are receiving considerable attention of reliability analysts. Mixtures of Weibull distributions have been considered by Jiang and Murthy (1995) in the context of modelling failure-time data. An important bivariate distribution — the bivariate Inverse Gaussian distribution, having wide applications in modelling bivariate failure-time data, has been considered by Kocherlakota (1986).

Mixture models, in general, are also important in reliability studies for “the overall failure distribution” of a multi-component item (Everitt, 1985, p. 560). When a product, say electric bulbs, is acquired in batches from two different machines or from the same machine on two different occasions, a mixture model seems to be appropriate. A contamination model is used when most of the items in the sample come from a known homogeneous population and a few from another unknown population (of the same family), possibly due to suspected lack of proper quality control in the process. In such a situation a reliability analyst may, at the initial stage, want to carry out a statistical test to ensure whether the items are really homogeneous, in order to avoid difficulties in dealing with a mixture distribution.

Now, in view of Lemma 7.2.1, it is clear that there exists a large class of reliability distributions and the associated class of mixture models for each of which characterization of an appropriate P^3 and the corresponding L-optimal test is possible. This can be stated formally in the following

Theorem 7.4.1. Consider a class $\mathcal{F}_0$ of reliability distributions satisfying conditions (C1) – (C5) of Lemma 7.2.1 and the corresponding class $\mathcal{G}_0$ of mixture probability distributions. Then each $g \in \mathcal{G}_0$ admits a P^3 , say η for which an L-optimal test is based on the statistic T obtained from the lemma. Furthermore, this test is consistent.

Proof. The first part follows from Lemma 7.2.1. Consistency of the test follows from the consistency of T as an estimator of η . $\square$

We now illustrate the above theorem through the following ten probability models $f^{(1)}, f^{(2)}, \dots, f^{(10)}$ and the corresponding mixture models $g^{(1)}, g^{(2)}, \dots, g^{(10)}$ which are very important and of wide use in reliability theory :

(i) Normal : $f^{(1)}(x|\theta) = (2\pi)^{-1/2} \exp \left\{ -\frac{1}{2}(x - \theta)^2 \right\}$; $\theta_0 = 0$.

(ii) Lognormal : $f^{(2)}(x|\theta) = (2\pi)^{-1/2} x^{-1} \exp \left\{ -\frac{1}{2}(\log x - \theta)^2 \right\}$; $\theta_0 = 0$.

(iii) Exponential : $f^{(3)}(x|\theta) = (1/\theta) \exp \{ -(x/\theta) \}$; $\theta_0 = 1$.

(iv) Inverse Gaussian with location :

$$f^{(4)}(x|\theta) = (2\pi x^3)^{-1/2} \exp \{ -(x - \theta)^2 / 2\theta^2 x \}$$
 ; $\theta_0 = 1$.

(v) Inverse Gaussian with dispersion :

$$f^{(5)}(x|\theta) = (\theta/2\pi x^3)^{1/2} \exp\{-(x - \theta)^2/2\theta^2 x\}; \theta_0 = 1.$$

(vi) Bivariate Exponential Conditional (BEC) :

$$f^{(6)}(x, y|\theta) = \beta\lambda v(\theta) \exp(\beta x + \lambda y + \theta\beta\lambda xy); \theta_0 = 0,$$

β, λ are known (each assumed to be equal to 1) (see Arnold and Strauss, 1988).

(vii) Geometric : $f^{(7)}(x|\theta) = \theta(1 - \theta)^{x-1}$; $\theta_0 = 1/2$.

(viii) Bivariate Inverse Gaussian:

$$\begin{aligned} f^{(8)}(x, y|\theta) = & \frac{1}{4\pi} \left\{ \frac{\lambda_1 \lambda_2}{x^3 y^3 (1 - \theta^2)} \right\}^{\frac{1}{2}} \left[\exp - \frac{1}{2(1 - \theta^2)} \right. \\ & \cdot \left\{ \frac{\lambda_1}{\mu_1^2} \frac{(x - \mu_1)^2}{x} - \frac{2\theta}{\mu_1 \mu_2} \left(\frac{\lambda_1 \lambda_2}{xy} \right)^{\frac{1}{2}} (x - \mu_1)(y - \mu_2) \right. \\ & \left. \left. + \frac{\lambda_2}{\mu_2^2} \frac{(y - \mu_2)^2}{y} \right\} + \exp - \frac{1}{2(1 - \theta^2)} \left\{ \frac{\lambda_1}{\mu_1^2} \frac{(x - \mu_1)^2}{x} \right. \right. \\ & \left. \left. + \frac{2\theta}{\mu_1 \mu_2} \left(\frac{\lambda_1 \lambda_2}{xy} \right)^{\frac{1}{2}} (x - \mu_1)(y - \mu_2) + \frac{\lambda_2}{\mu_2^2} \frac{(y - \mu_2)^2}{y} \right\} \right]; \end{aligned}$$

with $\lambda_1, \lambda_2, \mu_1, \mu_2$ all known and $\theta_0 = 0$ (see (1.1) of Kocherlakota, 1986).

(ix) Folded normal: $f^{(9)}(x|\theta) = \phi(x - \theta) + \phi(x + \theta)$; $\theta_0 = 0$, where $\phi(\cdot)$ is the density function of a standard normal variable. (See SenGupta and Pal, 1993a.)

(x) Weibull with shape parameter : $f^{(10)}(x|\theta) = \theta x^{\theta-1} \exp(-x^\theta)$; $\theta_0 = 1$.

In each of $f^{(6)}$ and $f^{(8)}$, θ plays the role of an 'interaction' parameter so that $\theta=0$ holds iff the components X and Y are independent. $f^{(7)}$ is appropriate as a model for the 'number of occasions' a component or a device is used before it fails to work.

Observe that the first seven density functions belong to the one parameter exponential family and using Corollary 7.2.1, one sees that the theorem

is applicable to the corresponding mixture models. We give the forms of η and T corresponding to each $g^{(i)}, i = 1, \dots, 7$:

$$\begin{aligned}\eta^{(1)} &= p\theta ; \quad T^{(1)} = \frac{1}{n} \sum_{i=1}^n X_i , \\ \eta^{(2)} &= p\theta ; \quad T^{(2)} = \frac{1}{n} \sum_{i=1}^n \log X_i , \\ \eta^{(3)} &= p(\theta - 1) ; \quad T^{(3)} = \frac{1}{n} \sum_{i=1}^n (X_i - 1) , \\ \eta^{(4)} &= p(\theta - 1) ; \quad T^{(4)} = \frac{1}{n} \sum_{i=1}^n (X_i - 1) , \\ \eta^{(5)} &= \frac{p}{2} \left(1 - \frac{1}{\theta}\right) ; \quad T^{(5)} = \frac{1}{2n} \sum_{i=1}^n \left(3 - X_i - \frac{1}{X_i}\right).\end{aligned}$$

For $g^{(6)}$, a little algebra gives

$$\begin{aligned}\eta^{(6)} &= 0 \text{ if } \theta = 0, \quad p \text{ arbitrary} \\ &= p \left\{ 1 - \frac{1 - v(\theta) + \theta}{\theta^2} \right\} \text{ if } \theta > 0, \quad p \text{ arbitrary.}\end{aligned}$$

Corresponding $T^{(6)} = \frac{1}{n} \sum_{i=1}^n (1 - X_i Y_i)$. Also

$$\eta^{(7)} = 2p(2\theta - 1)/\theta ; \quad T^{(7)} = (2/n) \sum_{i=1}^n (2 - X_i).$$

$f^{(8)}$ through $f^{(10)}$, though not members of the exponential family, may be seen to satisfy the conditions of Lemma 7.2.1. For $g^{(8)}$, let

$$U = \frac{\lambda_1(X - \mu_1)^2}{\mu_1^2 X}, \quad V = \frac{\lambda_2(Y - \mu_2)^2}{\mu_2^2 Y}.$$

Then the density of (U, V) is that of a bivariate chisquare distribution given by (2.3) of Kocherlakota (1986). So from Table 1 of Gunst and Webster (1973), it follows that,

$\eta^{(8)} = 2p\theta^2$ and $T^{(8)} = \frac{1}{n} \sum_{i=1}^n (U_i - 1)(V_i - 1)$, where (U_i, V_i) corresponds to (X_i, Y_i) .

For $g^{(9)}$, since the score function vanishes identically at $\theta_0 = 0$, we use a reparametrization $\theta' = \theta^2$ and note after a little algebra that

$$\eta^{(9)} = \frac{1}{2}p \left\{ I\left(\frac{\theta'}{2}, \frac{3}{2}\right) + \theta' \right\} \quad ; \quad T^{(9)} = \frac{1}{2n} \sum_{i=1}^n (X_i^2 - 1)$$

where

$$I(\lambda, r) = \frac{1}{\Gamma(r)} \int_0^\lambda \exp(-z) z^{r-1} dz.$$

Finally in case of $g^{(10)}$,

$$\left. \frac{\partial \log f^{(10)}(X|\theta)}{\partial \theta} \right|_{\theta_0} = 1 + \log X - X \log X.$$

Differentiating both sides of expression (7) of Johnson and Kotz (1970, p. 253) with respect to t and putting $t = 1$, one sees that $E_{f_\theta}(X \log X) = \frac{1}{\theta} \Gamma'(\frac{1}{\theta} + 1)$. Also $E_{f_\theta}(\log X) = -\gamma_E/\theta$, where γ_E is Euler's constant and hence one has

$$\eta^{(10)} = p \left[1 - \frac{\gamma_E}{\theta} - \frac{1}{\theta} \Gamma' \left(\frac{1}{\theta} + 1 \right) \right] \quad ; \quad T^{(10)} = \frac{1}{n} \sum_{i=1}^n (1 + \log X_i - X_i \log X_i).$$

Interestingly, though not obvious, it can be shown by differentiation that the expression for $\gamma(\theta, \theta_0)$ in this case is monotonically increasing in θ , so that (C5) is satisfied.

7.5. Estimating function approach

Estimating functions play an important role in statistical inference and have been used by a number of researchers. We shall see that this approach in mixtures may lead to asymptotically optimal tests in a certain class. For this we need a result which is derived below under a quite general setup.

Suppose X is a random variable whose distribution depends on ξ (the parameter of interest) and θ (the nuisance parameter, which may be vector-valued). Let $t(X)$ be a real-valued measurable function of X with $E(t(X)) = \xi$ and $\text{Var}(t(X)) = \sigma_{\xi, \theta}^2 < \infty$.

If $(X_1, X_2, \dots, X_n)$ is a random sample of size n from this distribution then $\sum_{i=1}^n l_{in}(t(X_i) - \xi)$ denotes a typical member of the class of linear unbiased estimating functions (LUEFs). Assume l_{in} 's satisfy

- (i) $\sum_{i=1}^n l_{in} = 1 \forall n$.
- (ii) $\frac{\left(\sum_{i=1}^n l_{in}^2\right)}{\max_{1 \leq i \leq n} l_{in}^2} \rightarrow \infty$ as $n \rightarrow \infty$.

Condition (ii) implies that $\sum_{i=1}^n l_{in}(t_i - \xi) \xrightarrow{\mathcal{L}} N(0, \sigma_{\xi, \theta}^2 \sum_{i=1}^n l_{in}^2)$ as $n \rightarrow \infty$ where $t_i = t(X_i)$ (Hajek and Sidak, 1967, p. 153). Suppose the null hypothesis to be tested is $H_0 : \xi = 0$. In two-component mixture models with unknown mixing proportion p , the role of the nuisance parameter θ is played by the unknown mixing proportion p . Under H_0 , $\sigma_{0, \theta}^2$ is usually free from θ . A reasonable test statistic for H_0 is

$$T_{ln} = \sum_{i=1}^n l_{in} t_i \quad (7.5.1)$$

which is obtained by putting $\xi = 0$ in the corresponding LUEF. Write $\bar{t}_n = \frac{1}{n} \sum_{i=1}^n t_i$. We now prove the following

Theorem 7.5.1. (a) For testing H_0 against the one-sided alternative $H_1 : \xi > 0$ ($\xi < 0$) the asymptotically level α test which rejects H_0 for large (small) values of $\bar{t}_n$ maximizes the asymptotic power uniformly at each $\xi > 0$ (< 0) and any fixed θ among all asymptotically level α tests of the class based on test-statistics of the form (7.5.1).

(b) For testing H_0 against the two-sided alternative $H_2 : \xi \neq 0$, the level α test which rejects H_0 for large values of $|\bar{t}_n|$ maximizes the asymptotic

power uniformly at each $\xi \neq 0$ and any fixed θ among all level α tests of the said class.

Proof. Let σ_0^2 be the value of $\sigma_{\xi,\theta}^2$ when H_0 is true. Note that under H_0 ,

$$T_{ln} \xrightarrow{\mathcal{L}} N \left(0, \sigma_0^2 \sum_{i=1}^n l_{in}^2 \right) \text{ as } n \rightarrow \infty.$$

Also under any ξ and θ

$$T_{ln} = \sum_{i=1}^n l_{in}(t_i - \xi) + \xi \sum_{i=1}^n l_{in} \xrightarrow{\mathcal{L}} N \left(\xi, \sigma_{\xi,\theta}^2 \sum_{i=1}^n l_{in}^2 \right) \text{ as } n \rightarrow \infty.$$

(a) If α be the asymptotic level and c_n be the cut-off point then

$$\alpha \doteq 1 - \Phi \left(\frac{c_n}{\sigma_0 \sum_i l_{in}} \right) \Rightarrow c_n \doteq \tau_\alpha \sigma_0 \left(\sum_i l_{in}^2 \right)^{1/2}.$$

Asymptotic power at any $\xi > 0$ and θ is

$$1 - \Phi \left(\frac{\tau_\alpha \sigma_0}{\sigma_{\xi,\theta}} - \frac{\xi}{\sigma_{\xi,\theta} (\sum_i l_{in}^2)^{1/2}} \right).$$

This will be maximum if $\sum_i l_{in}^2$ is minimum subject to (i). The maximum value is attained if $l_{in} = 1/n \ \forall n$. The same is also true when $\xi < 0$. This proves (a).

(b) In this case if d_n be the cut-off point, then

$$d_n \doteq \tau_{\alpha/2} \sigma_0 \left(\sum_i l_{in}^2 \right)^{1/2}.$$

The asymptotic power at any $\xi \neq 0$ is seen to be

$$\beta(u_{ln}) = 2 - \Phi \left(\frac{\tau_{\alpha/2} \sigma_0}{\sigma_{\xi,\theta}} - \frac{\xi}{\sigma_{\xi,\theta} u_{ln}} \right) - \Phi \left(\frac{\tau_{\alpha/2} \sigma_0}{\sigma_{\xi,\theta}} + \frac{\xi}{\sigma_{\xi,\theta} u_{ln}} \right)$$

where $u_{ln} = (\sum_i l_{in}^2)^{1/2}$. Hence

$$\frac{d\beta(u_{ln})}{du_{ln}} = \left[\phi \left(\frac{\tau_{\alpha/2} \sigma_0}{\sigma_{\xi,\theta}} + \frac{\xi}{\sigma_{\xi,\theta} u_{ln}} \right) - \phi \left(\frac{\tau_{\alpha/2} \sigma_0}{\sigma_{\xi,\theta}} - \frac{\xi}{\sigma_{\xi,\theta} u_{ln}} \right) \right] \left[\frac{\xi}{\sigma_{\xi,\theta} u_{ln}^2} \right].$$

Observe that if $\xi > 0$, the first factor of the above expression is negative whereas the second factor is positive. Also if $\xi < 0$, the first factor is positive while the second one is negative. In either case we have $\frac{d\beta(u_{ln})}{du_{ln}} < 0$. Hence $\beta(u_{ln})$ is decreasing in u_{ln} . Consequently the maximum value of the asymptotic power is attained when u_{ln} is the minimum, *i.e.*, when $l_{in} = 1/n$ for all i . This proves (b). Hence the theorem. $\square$

Let us now consider the mixture model (7.2.1) whose component density $f(x|\theta)$ satisfies (A1) - (A5) of Section 7.2. Take t_i to be the score function based on X_i . Then, after a renaming of the parameters, we can have $\xi = pE_{f_{\theta_0}} \left(\frac{\partial}{\partial \theta} \log f(x|\theta) \Big|_{\theta_0} \right)$ as the parameter of interest and p as the nuisance parameter. Thus ξ is nothing but the appropriate P^3 . Appealing to Theorem 7.5.1 we can then have the following

Corollary 7.5.1. The test based on the average score statistic, as given in Theorem 7.2.1, is optimal in the sense given in Theorem 7.5.1, among all tests based on LUEFs.

7.6. Examples

In this Section, we illustrate the applications of the L-optimal test considered in Section 4 with the help of two data sets — one on rock specimens and the other on a simulated sample from a mixture distribution. In both the examples we see how elegantly the test can be applied to detect the presence of possible contamination in the population.

7.6.1. Rock data

Barnett and Lewis(1994, p. 253, Example 6.4) present the data on Sulphate Soundness Test (SST) values (X) for $n = 12$ rock specimens originally considered by the Department of Main Roads, New South Wales, Australia. Further to the observations made from their empirical findings we noted that the distribution of X can be reasonably modelled by a single exponential or a mixture of two exponentials. Since the data show a sharp dominance of the ordered observations from the seventh onwards over the first six, it is more reasonable to assume a mixture of two exponentials as

our underlying model. Note that if Z follows exponential distribution with mean 1, then $Pr\{Z > 6.5\} = .0015$, so that only .15% of the observations, drawn from such a population, exceed 6.5 on an average. However about 50% of the observations in the sample at hand exceed 6.5. This fact also suggests that there is a distinct shift in mean from 1 (of one component density) to some unknown θ (of the other component from which the last six ordered observations constitute a sample). A plausible mixture model is, therefore, given by $g^{(3)}$ with $\theta_0 = 1$. The L-optimal test statistic is equivalent to $T = \sum_{i=1}^n X_i$ and $2T$ follows χ^2 distribution with $2n$ d.f. under H_0 . Since $2T = 204$ and $\chi_{.01,24}^2 = 42.98$, the data are strongly in support of a possible contamination, as expected.

Let us now carry out a goodness-of-fit test, as a post-analysis, to judge the plausibility of $g^{(3)}$ as a model for these data. The moment estimates of p and θ are .49 and 16.2 respectively. Assuming these as the true values of the parameters, the value of the Kolmogorov-Smirnov D_n statistic comes out as .3044, which is insignificant as compared to the 5% critical value of .3754. We therefore reach the same conclusion as the foregoing one about the presence of a contamination, further validating our simple P^3 approach.

7.6.2. Simulated lognormal mixture data

The following $n = 20$ observations are drawn from a mixture of two lognormal distributions given by the density $g^{(2)}$ with $\theta = 0.5$ and $p = 0.7$:

0.525	3.897	2.551	1.010	3.970	1.029	1.980	0.120	0.423	0.431
3.259	0.127	0.877	2.625	0.693	1.359	2.154	2.764	2.869	1.801

Observations from a lognormal distribution have been generated using the RNLNL subroutine of IMSL. Uniform random variables used for the purpose of 'mixing' have been generated by the RNUN subroutine of the same library.

Here $\tau = \sqrt{n}T^{(2)} = (1/\sqrt{n}) \sum_{i=1}^n \log X_i \sim N(0, 1)$ and observed $\tau = 0.7263$, which is insignificant at 5% level. Hence the test accepts the null hypothesis $H'_0 : \eta^{(2)} = 0$ against either of the one-sided alternatives at the said level.

7.7. L-optimal tests in the presence of nuisance parameters

Let us now consider the situation where apart from the unknown mixing proportion p , some nuisance parameters are also present in the mixture model. In such a situation, as indicated before, we can conceive of L-optimal similar or L-optimal invariant tests for appropriately chosen P^3 s. Even when no small sample L-optimal test can be constructed, we can think of L-optimal $C(\alpha)$ tests, *i.e.*, tests which are optimal $C(\alpha)$ against an appropriate sequence of alternatives for each given p .

Let

$$g(x|p, \vartheta, \theta) = pf(x|\vartheta, \theta) + qf(x|\vartheta, \theta_0) \quad (7.7.1)$$

be a mixture density in which θ is the unknown parameter of interest in Θ , θ_0 is a known point of Θ , ϑ is the unknown nuisance parameter (may be vector-valued), p is the unknown mixing proportion and f is a density with respect to some σ -finite measure. The following definitions, which are generalizations of the earlier ones under the presence of nuisance parameters, are relevant in this context :

Definition 7.7.1. A parametric function $\eta = \eta(p, \vartheta, \theta, \theta_0)$ is said to be a P^3 if $\eta = 0$ iff $p = 0$ or $\theta = \theta_0$ or $p = 0, \theta = \theta_0$, *i.e.*, iff 'no mixture' holds for all ϑ .

Definition 7.7.2. A (randomized) test ψ based on an unbiased and/or consistent estimator T of η is said to be L-optimal similar unbiased (LSU) or L-optimal invariant unbiased (LIU) or L-optimal $C(\alpha)$ (L- $C(\alpha)$) for testing $H_0 : \eta = 0$ against $H_1 : \eta \neq 0$ according as ψ is LMPSU or LMPIU or optimal $C(\alpha)$ for testing $H_0 : \theta = \theta_0$ against $\theta \neq \theta_0$ for each given $p \in (0, 1)$.

We illustrate the above constructions of the P^3 s and the corresponding L-optimal tests by means of three examples considered in Chapter 2. The models are

- (i) $g(x|p, \mu_1, \mu_2) = p\phi(x - \mu_1) + (1 - p)\phi(x - \mu_2)$
- (ii) $g(x|p, \mu, \sigma^2) = \frac{p}{\sigma}\phi\left(\frac{x - \mu}{\sigma}\right) + \frac{1 - p}{\sigma}\phi\left(\frac{x}{\sigma}\right)$
- (iii) $g(x|p, \mu_1, \mu_2, \sigma^2) = \frac{p}{\sigma}\phi\left(\frac{x - \mu_1}{\sigma}\right) + \frac{1 - p}{\sigma}\phi\left(\frac{x - \mu_2}{\sigma}\right)$

In each of the above models we want to test H_0 : 'no mixture' against the alternative H_1 : 'there is mixture' on the basis of n i.i.d. observations. If p is given and belongs to $(0, 1)$ we see in Chapter 2 that after suitable reparametrizations in the first two models the problem can be reduced by location invariance and scale invariance respectively. The LMPSU and LMPIU tests for the parameters of interest against two-sided alternatives have been derived in Chapter 2 and are based, respectively, on the critical regions

$$\begin{aligned}\omega_1 &= \{(X_1, X_2, \dots, X_n) : s^2 > c\} \\ \omega_2 &= \{(X_1, X_2, \dots, X_n) : \bar{X}^2/s^2 > c\}\end{aligned}$$

where $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ and $s^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$.

Let us now consider model (iii). Recall that for given p this model (or a reparametrized version of it), as considered in Chapter 2, does not admit an unconditional LMPSU test for testing the hypothesis of no mixture against two-sided alternatives. Also the LMPIU test is not unique. We have derived an optimal third order $C(\alpha)$ test in Chapter 6. This test has the critical region

$$\omega_3 = \left\{ (X_1, X_2, \dots, X_n) : \frac{\left\{ \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^3 \right\}^2}{s^6} > c \right\}.$$

In the light of the definitions (7.7.1) and (7.7.2), the forms of the appropriate P^3 s and the corresponding unbiased and/or consistent estimators for each of the models (i), (ii) and (iii) are at once clear. These are, respectively,

$$\begin{aligned}\eta_1 &= p(1-p)(\mu_1 - \mu_2)^2; \quad T_1 = s^2 - 1, \\ \eta_2 &= \frac{p^2 \mu^2}{p(1-p)\mu^2 + \sigma^2}; \quad T_2 = \bar{X}^2/s^2 \quad \text{and} \\ \eta_3 &= \frac{\{p(1-p)(2p-1)\}^2 \Delta^6}{(1+p(1-p)\Delta^2)^3}; \quad T_3 = \frac{\left\{ \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^3 \right\}^2}{s^6},\end{aligned}$$

where $\Delta = (\mu_1 - \mu_2)/\sigma$. Again with reference to the reparametrized models (i'), (ii') and (iii') (see Chapter 2), these P^3 s take the form

$$\eta_1 = p(1-p)\theta^2,$$

$$\eta_2 = \frac{p^2 \theta^2}{p(1-p)\theta^2 + 1}$$

and

$$\eta_3 = \frac{\{p(1-p)(2p-1)\}^2 \theta^6}{(1+p(1-p)\theta^2)^3}$$

respectively, which are seen to be unrelated to the nuisance parameters. Recall that the above expression for η_3 as a measure of population skewness for a mixture of two univariate normal distributions follows directly from (2.14) of Mardia (1974).

Combining all these, we have the following

Theorem 7.7.1. (a) If the unknown $p \in [0, 1)$, the critical region ω_1 is LSU for testing $H_0 : \eta_1 = 0$ against $H_1 : \eta_1 \neq 0$ in model (i).

(b) If the unknown $p \in [0, 1)$, the critical region ω_2 is LIU for testing $H_0 : \eta_2 = 0$ against $H_1 : \eta_2 \neq 0$ in model (ii).

(c) If the unknown $p \in [0, 1) \setminus \{\frac{1}{2}\}$, the critical region ω_3 is L-optimal third order $C(\alpha)$ for testing $H_0 : \eta_3 = 0$ against $H_1 : \eta_3 \neq 0$ in model (iii).

The test given by the critical region ω_1 for model (i) has the interesting properties of unbiasedness and consistency as are seen from the following

Theorem 7.7.2. The test given by the critical region ω_1 is unbiased and consistent.

Proof. Following Behboodian (1972) and using the same notations one can write

$$P_{\eta_1} \left\{ \sum_{i=1}^n (X_i - \bar{X})^2 > C \right\} = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} P_{\eta_1} \{T_k > C\} \quad (7.7.2)$$

where $T_k = kS_{k1}^2 + (n-k)S_{k2}^2 + k(n-k)(\bar{X}_{k1} - \bar{X}_{k2})^2/n$; $k = 2, 3, \dots, n-2$ and T_k is defined in an obvious way for $k = 0, 1, n-1$ and n .

It can be easily seen that

$$kS_{k1}^2 \equiv \chi_{(1)}^2 \sim \chi_{k-1}^2, \quad (n-k)S_{k2}^2 \equiv \chi_{(2)}^2 \sim \chi_{n-k-1}^2$$

and

$$k(n-k)(\bar{X}_{k1} - \bar{X}_{k2})^2/n \equiv \chi_{(3)}^2 \sim \chi_{1, \epsilon^2}^2,$$

δ^2 being the noncentrality parameter $(\mu_1 - \mu_2)^2$. Further all these variables are independent. Obvious modifications are to be made for $k = 0, 1, n - 1$ and n . Recalling that the tail probability of a non-central chisquare variable is an increasing function of its noncentrality parameter (Result C1 of Appendix C), it follows from (7.7.2) that for any $\eta_1 > 0$

$$\begin{aligned} P_{\eta_1} \left\{ \sum_{i=1}^n (X_i - \bar{X})^2 > C \right\} &> \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} P_{\eta_1} \left\{ \chi_{(1)}^2 + \chi_{(2)}^2 + \chi_{(3)}^{*2} > C \right\} \\ &> P_0 \left\{ \chi_{n-1}^2 > C \right\} = \alpha \end{aligned} \quad (7.7.3)$$

where $\chi_{(3)}^{*2}$ is distributed as a χ_{1, η_1}^2 variable. This shows that the above test is unbiased.

To prove consistency we see by CLT (Cramer, 1946, p. 365) that the asymptotic power at any η_1 is

$$1 - \Phi \left(\sqrt{n}(C^* - \eta_1 - 1)/(\gamma_2 + 2)^{\frac{1}{2}} \right) \quad (7.7.4)$$

where $C^* = (\sqrt{2(n-1)\tau_\alpha} + n - 1)/n$ is the asymptotic expression for the cut-off point and γ_2 is the usual measure of kurtosis for this population. The expression in (7.7.4) tends to 1 as $n \rightarrow \infty$ implying the consistency of the test. Hence the theorem. $\square$

Remark 7.7.1. Consider the SWS-CU mixture family

$$g(x|a, \rho, \mu_0, p) = \frac{p}{2\pi} \left\{ 1 + 2 \sum_{r=1}^{\infty} \rho^r \cos r(x - \mu_0) \right\} + \frac{1-p}{2\pi},$$

the discussion on the P^3 aspect of which was abandoned in Chapter 4.

Observe that

$$E_g(\cos(X_i - X_j)) = (p\rho)^2 \equiv \eta, \text{ say}$$

where η plays the role of the appropriate P^3 in this case. An unbiased and consistent estimator of η is (SenGupta and Pal, 2001)

$$U = \frac{2}{n(n-1)} \sum_{i < j} \cos(X_i - X_j) = \frac{R^2}{n(n-1)} - \frac{1}{n-1}.$$

A test for isotropy (which is now equivalent to $H'_0 : \eta = 0$) may be based on U or equivalently on R^2 . Chang (1991) has shown that for $p = 1$ and unknown μ_0 , the LMPI test for $H_0 : \rho = 0$ against $H : \rho > 0$ is the Rayleigh's test, *i.e.*, one having the critical region

$$\omega = \{(X_1, X_2, \dots, X_n) : R^2 > c\}.$$

Note that this test does not depend on a , and is therefore robust optimal against the SWS family.

Remark 7.7.2. Mixtures of SMN distributions (SenGupta and Pal, 1991; 1993b) deserve mention in this context.

Let

$$\begin{aligned} g_{(p)}(x|m, \sigma^2, \rho) &= p(2\pi\sigma^2)^{-k/2} |\Sigma_\rho|^{-1/2} \exp \left\{ -(x - M)' \Sigma_\rho^{-1} (x - M) / 2\sigma^2 \right\} \\ &+ q(2\pi\sigma^2)^{-k/2} \exp \left\{ -(x - M)' (x - M) / 2\sigma^2 \right\} \end{aligned}$$

be the density of a mixture of two k -variate SMN distributions with the equicorrelation coefficient ρ of the correlation matrix Σ_ρ of the second component having the value 0 and other symbols having their usual meanings. For testing $H_0 : \rho = 0$ against $H_1 : \rho > 0$ when p is known, we considered in Chapter 5 the LMP and LMPS tests under different situations regarding the knowledge of the other parameters μ and σ^2 . Four cases *viz.* m, σ^2 both known (Case 1), m known, σ^2 unknown (Case 2), m, σ^2 both unknown (Case 3) and m unknown, σ^2 known (Case 4) were treated separately and locally optimal tests (LMP in Case 1 and LMPS in Cases 2-4), based on the test statistics T_1 through T_4 , were constructed. Explicit forms of the T_i s are given in SenGupta and Pal (1993b).

Now suppose p also is unknown. One can verify following some simple derivations that $T_1/nk(k-1)$, T_2-1 , $T_3/(1-T_3)-1/(k-1)$ and T_4/n are consistent estimators of $\eta_1 = p\rho$, $\eta_2 = (k-1)p\rho$, $\eta_3 = kp\rho/(k-1)(1-p\rho)$ and $\eta_4 = k(k-1)p\rho$ respectively, where the η_i s serve as the appropriate P^3 s in the respective cases. Further each of these estimators is equivalent to the corresponding test statistic. Hence the locally optimal tests in all the cases are also L-optimal or L-optimal similar for testing the hypotheses of 'no mixture' expressed in terms of the null values of the η_i s.

Chapter 8

Locally Bahadur Optimal Tests in Mixtures

8.1. Introduction

In the previous Chapters so far, we confined ourselves mainly to tests within the framework of Neyman-Pearson theory, except in the situations where we adopted the $C(\alpha)$ approach or the approach of Davies. Now we take a different view-point. We attempt to propose tests for ‘no contamination’ which are optimal in terms of having the largest slope locally in the sense of Bahadur (1960; 1967). Such a test is called locally Bahadur optimal.

Bahadur optimality of a test, or more specifically of a sequence of test statistics, is a very important and useful optimality criterion in large samples. The asymptotic performance of a test is judged in terms of its *exact slope* and one having the largest slope for all values of the parameter under the alternative hypothesis is called Bahadur optimal. It has been shown by Bahadur (1967) that LRTs are optimal in this sense under certain regularity conditions – although other tests with such an optimality property do exist within a restricted class. In the absence of tests which are (globally) Bahadur optimal, one may search for a locally optimal test. Also in many cases, global optimality of a given test is difficult to establish while local optimality can be established rather easily. Moreover, the class of locally Bahadur optimal tests is usually wider than the class of globally optimal tests.

In this Chapter we first present the concept of local Bahadur optimality from Bahadur (1960, 1967) in Section 8.2. In the next Section we consider a two-component mixture of distributions from the one-parameter exponential family and propose a test for the parameter against two-sided alternatives, assuming the mixing proportion to be known. Note that for such a

testing problem, the LRT is not available in a closed form, neither its Bahadur optimality can be verified easily. Also the LMPU test, as considered by Durairajan (1980) under a general setting, is quite complicated and its Bahadur optimality or otherwise is not known. We show that the proposed test is locally Bahadur optimal by verifying certain sufficient conditions of Hsieh (1979). We also present some examples of mixture distributions (Section 8.4) for which the conditions are satisfied. This Chapter is based on Pal (2000a).

8.2. Local Bahadur optimality of a test

Let the null hypothesis be $H_0 : \theta \in \Theta_0$ which is to be tested against the alternative $H_1 : \theta \in \Theta_1 = \Theta - \Theta_0$ where θ is an unknown parameter. Let $\{T^{(n)}\}$ be a sequence of test statistics such that large values of $T^{(n)}$ indicate significance. For any t and θ , let

$$F_n(t, \theta) = P_\theta(T^{(n)} < t)$$

and

$$G_n(t) = \inf \{F_n(t, \theta) : \theta \in \Theta_0\}.$$

Write $L_n = 1 - G_n(T^{(n)})$. If there exists a function $c(\theta)$ defined over Θ_1 such that $n^{-1} \log L_n \rightarrow -\frac{1}{2}c(\theta)$ as $n \rightarrow \infty$ a.s. when θ obtains, then $c(\theta)$ is called the *exact slope* of the sequence $\{T^{(n)}\}$. Suppose $f(x, \theta)$ is the density function of a single observation X with respect to some σ -finite measure μ . For any θ and θ_0 , let the Kullback-Leibler information be defined by

$$K(\theta, \theta_0) = E_{f_\theta} \left[\log \frac{f(x, \theta)}{f(x, \theta_0)} \right].$$

Denote $J(\theta) = \inf \{K(\theta, \theta_0) : \theta_0 \in \Theta_0\}$. Bahadur (1967) showed that $c(\theta) \leq 2J(\theta)$ for each non-null θ . A sequence $\{T^{(n)}\}$ for which $c(\theta) = 2J(\theta)$ for all non-null θ is called Bahadur optimal.

A sequence of test statistics $\{T^{(n)}\}$ with exact slope $c(\theta)$ is said to be locally optimal (Bahadur, 1967) if for each $\theta_0 \in \Theta_0$, $c(\theta) \sim 2J(\theta)$ as $\theta \rightarrow \theta_0$. The class of locally optimal sequences, as pointed out by Bahadur, is much wider than the class of sequences having exact slope $2J(\theta)$ for each non-null θ , especially if Θ_0 consists of a single point θ_0 .

8.3. A locally optimal test in mixtures from the one-parameter exponential family

Consider the mixture model

$$g(x|p, \theta) = pf(x|\theta) + (1 - p)f(x|\theta_0) \quad (8.3.1)$$

where $f(x|\theta)$ is the density (w.r.t. some σ -finite measure μ) of the one-parameter exponential family of distributions given by

$$f(x|\theta) = B(\theta) \exp [\psi(\theta)U(x)] h(x). \quad (8.3.2)$$

Here $\theta \in \Theta$, a non-degenerate interval of the real line and $x \in \mathcal{X}$. Let us make the following assumptions:

(A1) $f(x|\theta)$ is in the mean-value parametrization, so that $E_{f_\theta}(U(X)) = \theta \forall \theta \in \Theta$.

(A2) $\psi(\theta)$ is increasing in θ .

In (8.3.1), θ_0 is a known point of Θ and p ($0 < p < 1$) is assumed to be known. We want to test $H_0 : \theta = \theta_0$ against the alternative $H_1 : \theta \neq \theta_0$. Suppose $\{X_1, X_2, \dots\}$ is a sequence of *iid* observations drawn from (8.3.1). For the first n observations, define $\bar{U}_n = \frac{1}{n} \sum_{i=1}^n U(X_i)$. The sequence of test statistics we propose for testing H_0 against H_1 is $\{T^{(n)}\}$, where

$$T^{(n)} = \log \frac{B(\bar{U}_n)}{B(\theta_0)} + \bar{U}_n (\psi(\bar{U}_n) - \psi(\theta_0)). \quad (8.3.3)$$

It can be verified using (A2) that $T^{(n)} \equiv T^{(n)}(\bar{U}_n)$ is a convex function of $\bar{U}_n$ and since large or small values of $\bar{U}_n$ indicate departure from H_0 , we reject H_0 for large values of $T^{(n)}$. The above sequence is locally optimal as is seen from Theorem 8.3.1 below. To this end, recall expressions (2.6) and (2.9) of Hsieh (1979) which are sufficient conditions for establishing (global) Bahadur optimality of any sequence of test statistics. These conditions are, respectively,

(i) for each θ in $\Theta - \Theta_0$,

$$\liminf_{n \rightarrow \infty} T^{(n)} \geq J(\theta) \text{ a.e. } [P_\theta].$$

and

$$(ii) \quad \limsup_{n \rightarrow \infty} n^{-1} \log P_{\theta_0} [T^{(n)} \geq t] \leq -t \text{ for any } t > 0 \text{ and } \theta_0 \in \Theta_0.$$

(i) and (ii) are again ‘refined’ versions of conditions 1 and 2 of Bahadur and Raghavachari (1972). We prove the theorem by establishing a lemma under an additional assumption (A3) which ensures (ii) and then verifying (i) in the limit as $\theta \rightarrow \theta_0$.

(A3) The density $f(x|\theta)$ is such that

$$\limsup_{n \rightarrow \infty} E_{f_{\theta_0}} \left[\left\{ \frac{B(\bar{U}_n)}{B(\theta_0)} \right\}^{nh} \exp \left\{ nh \bar{U}_n (\psi(\bar{U}_n) - \psi(\theta_0)) \right\} \right] < \infty$$

for all $h \in (0, 1)$.

Lemma 8.3.1. Under (A3), for any $t > 0$

$$\limsup_{n \rightarrow \infty} n^{-1} \log P_{f_{\theta_0}} [T^{(n)} \geq t] \leq -t.$$

Proof. Proceeding as in Hsieh (1979) one sees by Markov’s inequality that for any $t > 0$ and for all $h \in (0, 1)$,

$$\begin{aligned} P_{f_{\theta_0}} [T^{(n)} \geq t] &= P_{f_{\theta_0}} [e^{nhT^{(n)}} \geq e^{nht}] \\ &\leq e^{-nht} E_{f_{\theta_0}} [e^{nhT^{(n)}}] \\ &= e^{-nht} E_{f_{\theta_0}} \left[\left\{ \frac{B(\bar{U}_n)}{B(\theta_0)} \right\}^{nh} \exp \left\{ nh \bar{U}_n (\psi(\bar{U}_n) - \psi(\theta_0)) \right\} \right], \end{aligned}$$

so that

$$\begin{aligned} n^{-1} \log P_{f_{\theta_0}} [T^{(n)} \geq t] &\leq -ht \\ &\quad + n^{-1} \log E_{f_{\theta_0}} \left[\left\{ \frac{B(\bar{U}_n)}{B(\theta_0)} \right\}^{nh} \exp \left\{ nh \bar{U}_n (\psi(\bar{U}_n) - \psi(\theta_0)) \right\} \right]. \end{aligned}$$

Hence

$$\limsup_{n \rightarrow \infty} n^{-1} \log P_{f_{\theta_0}} [T^{(n)} \geq t] \leq -ht \quad \text{by (A3) .}$$

Since this is true for all $h \in (0, 1)$, the lemma follows. $\square$

Theorem 8.3.1. If (A3) holds the sequence $\{T^{(n)}\}$ given by (8.3.3) is locally optimal for testing H_0 against H_1 in the sense of exact Bahadur slope.

Proof. That (ii) holds is immediate from the lemma. (i) is also seen to hold in the limit as $\theta \rightarrow \theta_0$, since

$$\lim_{n \rightarrow \infty} T^{(n)} = \log \frac{B(p\theta + q\theta_0)}{B(\theta_0)} + (p\theta + q\theta_0)(\psi(p\theta + q\theta_0) - \psi(\theta_0)) \quad \text{a.s. } [P_\theta]$$

by (A1) and the strong law of large numbers, where $q = 1 - p$ and P_θ is the probability measure dominated by μ under the mixture distribution (8.3.1). Clearly the right side tends to 0 as $\theta \rightarrow \theta_0$. Also $J(\theta) \rightarrow 0$ as $\theta \rightarrow \theta_0$. $\square$

Remark 8.3.1. Since $T^{(n)}$ is convex in $\bar{U}_n$, the critical region $\{T^{(n)} > t^{(n)}\}$ is equivalent to $\{\bar{U}_n > u_{1n}\} \cup \{\bar{U}_n < u_{2n}\}$ where u_{1n} and u_{2n} satisfy

$$\log B(u_{1n}) + u_{1n}(\psi(u_{1n}) - \psi(\theta_0)) = \log B(u_{2n}) + u_{2n}(\psi(u_{2n}) - \psi(\theta_0)). \quad (8.3.4)$$

(8.3.4), along with the size condition, may be used to find the cut-off points u_{1n} and u_{2n} of the test. The form of the LMPU test in mixtures of distributions from the exponential family does not, in general, coincide with our proposed test, as can be seen from the example of normal mixtures considered by Durairajan (1980). Our procedure thus provides a *new* test, which, apart from being simple, enjoys the property of local Bahadur optimality.

Remark 8.3.2. In the non-mixture case, *i.e.*, for $p = 1$ in (8.3.1), the test given by the statistic $T^{(n)}$ is equivalent to the likelihood ratio test, as can be verified easily. Equality in (i) is seen to hold in this case for all $\theta \in \Theta_1$. Again the lemma above ensures that (ii) also holds. Hence the proposed test is globally Bahadur optimal for testing H_0 against H_1 in the non-mixture model (8.3.2).

8.4. Examples

We now consider some specific distributions from the exponential family satisfying (A1) and (A2) and verify that for each of these models (A3) also holds. It then follows that in each of the examples, the sequence of tests as given by $T^{(n)}$ of (8.3.3) is locally Bahadur optimal.

Example 8.4.1. *Normal with known variance.* Let $f(x|\theta)$ be the density function of a normal distribution with mean θ and variance unity. Take $\theta_0 = 0$. Here

$$\left\{ \frac{B(\bar{U}_n)}{B(\theta_0)} \right\}^{nh} \exp \{ nh \bar{U}_n (\psi(\bar{U}_n) - \psi(\theta_0)) \} = \exp(nh \bar{X}_n^2/2),$$

where $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$, so that

$$E_{f_{\theta_0}} [\exp(nh \bar{X}_n^2/2)] = (1-h)^{-1/2} < \infty \quad \forall n.$$

Example 8.4.2. *Normal with known mean.* Let $f(x|\theta)$ be the density of a normal distribution with mean 0 and variance θ^2 , so that $B(\theta) = \theta^{-1/2}$, $\psi(\theta) = -1/2\theta$ and $U(X) = X^2$. Take $\theta_0 = 1$. Writing Y for $\bar{U}_n$, we see that

$$\left\{ \frac{B(Y)}{B(\theta_0)} \right\}^{nh} \exp \{ nh Y (\psi(Y) - \psi(\theta_0)) \} = Y^{-nh/2} \exp(-nh/2) \exp(nhY/2).$$

Hence, noting that $nY \sim \chi_n^2$ under H_0 , one can write

$$\begin{aligned} & E_{f_{\theta_0}} \left[\left\{ \frac{B(Y)}{B(\theta_0)} \right\}^{nh} \exp \{ nh Y (\psi(Y) - \psi(\theta_0)) \} \right] \\ &= \exp(-nh/2) \frac{n^{n/2}}{2^{n/2} \Gamma(n/2)} \int_0^\infty \exp \left(-\frac{n(1-h)}{2} y \right) y^{\frac{n(1-h)}{2}-1} dy \quad (8.4.1) \\ &= \exp(-nh/2) \cdot \frac{n^{n/2}}{2^{n/2} \Gamma(n/2)} \cdot \frac{\Gamma\{n(1-h)/2\}}{\{n(1-h)/2\}^{n(1-h)/2}}. \end{aligned}$$

Letting $m = n/2$ and using Stirling's approximation for gamma functions, one can see that the last term of (8.4.1) tends to $(1-h)^{-1/2}$ as $m \rightarrow \infty$.

Example 8.4.3. *Exponential with mean θ .* Consider now an exponential distribution with density $f(x|\theta) = (1/\theta) \exp(-x/\theta)$, with $\theta_0 = 1$. Proceeding as in Example 8.4.2, It can be verified after a little algebra that the desired limit in (A3) is again $(1-h)^{-1/2}$.

Example 8.4.4. *Poisson with mean θ .* Let us now consider a Poisson distribution with mean θ . Take $\theta_0 = 1$. Here $B(\theta) = \exp(-\theta)$, $\psi(\theta) = \log \theta$ and $U(X) = X$. One can then write

$$\begin{aligned} & E_{f_{\theta_0}} \left[\left\{ \frac{B(\bar{U}_n)}{B(\theta_0)} \right\}^{nh} \exp \{nh\bar{U}_n(\psi(\bar{U}_n) - \psi(\theta_0))\} \right] \\ &= \exp(-n(1-h)) + \exp(-n(1-h)) \sum_{z=1}^{\infty} (\exp(-z)z^z)^h \cdot \frac{n^{z(1-h)}}{z!}, \end{aligned}$$

where 0^0 is taken to be 1 by convention.

For any fixed $h \in (0, 1)$, define

$$\xi_z(n) = \{\exp(-z)z^z\}^h \exp\{-n(1-h)\}n^{z(1-h)}/z!; \quad n, z \in \mathbf{I}^+,$$

where $\mathbf{I}^+$ is the set of all positive integers.

The series $\sum_{z=1}^{\infty} \xi_z(n)$, for any fixed n , is seen to converge by the root test.

Also $\lim_{n \rightarrow \infty} \xi_z(n) = 0$ ($z \in \mathbf{I}^+$). Further

$$\begin{aligned} \sup_{z \in \mathbf{I}^+} \xi_z(n) &\leq \sup_{z \in \mathbf{I}^+} \{\exp(-n)n^z/z!\}^{1-h} \\ &= \{\exp(-n)n^n/n!\}^{1-h} \\ &\doteq (2\pi)^{-\frac{1-h}{2}} n^{-\frac{1-h}{2}} \rightarrow 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

which implies that

$$\lim_{n \rightarrow \infty} \{\sup_{z \in \mathbf{I}^+} \xi_z(n)\} = 0$$

and hence $\lim_{n \rightarrow \infty} \xi_z(n) = 0$ uniformly in $z \in \mathbf{I}^+$. These together imply that

$\limsup_{n \rightarrow \infty} \sum_{z=1}^{\infty} \xi_z(n) < \infty$, as is seen below:

Fix any $\epsilon > 0$. For any $n \in \mathbf{I}^+$, let $z_0 = z_0(\epsilon, n)$ be the smallest integer such that $\sum_{z=z_0+1}^{\infty} \xi_z(n) \leq \epsilon$. Again $\exists n_0 = n_0(\epsilon)$ (not depending on z) such that

$$\xi_z(n) \leq \epsilon \quad \forall n > n_0 \quad \text{whatever } z \in \mathbf{I}^+ \text{ may be.}$$

Hence whenever $n > n_0$

$$\begin{aligned}\sum_{z=1}^{\infty} \xi_z(n) &= \sum_{z=1}^{z_0} \xi_z(n) + \sum_{z=z_0+1}^{\infty} \xi_z(n) \\ &\leq \epsilon(z_0 + 1) \\ \Rightarrow \sup_{n > n_0} \sum_{z=1}^{\infty} \xi_z(n) &\leq \epsilon \left\{ \sup_{n > n_0} z_0(\epsilon, n) + 1 \right\}.\end{aligned}$$

Since $\xi_z(n) \rightarrow 0$ as $n \rightarrow \infty$, $z_0(\epsilon, n)$ is bounded for $n > n_0$. So

$$\sup_{n > n_0} \sum_{z=1}^{\infty} \xi_z(n) \leq k \text{ for some } k > 0.$$

Take

$$M = \max \left\{ \sum_{z=1}^{\infty} \xi_z(1), \sum_{z=1}^{\infty} \xi_z(2), \dots, \sum_{z=1}^{\infty} \xi_z(n_0), k \right\}.$$

It then follows that $\sum_{z=1}^{\infty} \xi_z(n) \leq M$ for all $n \in \mathbb{I}^+$ and hence

$$\limsup_{n \rightarrow \infty} \sum_{z=1}^{\infty} \xi_z(n) \leq M < \infty.$$

Chapter 9

Tests for Multiparameter Hypotheses in Possibly Multicomponent Mixtures

9.1. Introduction

In this Chapter we consider the problem of testing multiparameter hypotheses in mixtures of distributions when the number of components in the mixture may possibly be more than two. The problem of testing a multiparameter hypothesis, in a general setting, is quite old and at the same time important because in practical situations one comes across distributions specified by a vector parameter rather than by a single parameter. In many cases the LRT or Rao's test based on efficient scores can be conveniently derived. But in mixtures the derivations are often quite involved. Also 'their small sample performances relative to reasonable competitors, however, are nearly always unknown' (SenGupta and Vermeire, 1986, henceforth SG-V). There are alternative small sample procedures which are based on certain local optimality considerations, *e.g.*, Type C test of Neyman and Pearson (1938) that maximizes the local power in a predetermined direction, locally most total powerful unbiased (LMTPU) test of Isaacson (1951) or Type D and Type E (in the presence of nuisance parameters) tests that maximize the total curvature of the power function at the null hypothesis (Lehmann, 1986, p. 529). SG-V developed locally most mean power unbiased (LMMPU) tests which lead the above classical tests in many theoretical and computational aspects and can be obtained by directly appealing to the generalized Neyman-Pearson lemma.

The present Chapter is organized in the following way. In Section 9.2 we present some relevant definitions and theorems, from SG-V, in connection with the existence and construction of LMMPU tests. In Section 9.3 we

derive the general form of the LMMPU test for the hypothesis of no mixture, formulated as a multiparameter hypothesis, in a $(k + 1)$ -component mixture of univariate one-parameter distributions. The result is then applied to the case when the component distributions are normal. The case of two-component multivariate mixtures with known mixing proportion p is treated in Section 9.4.

In Section 9.5 we consider two-component mixtures of multivariate distributions where p is also unknown. We take up, as examples, the multivariate analogs of the univariate normal models discussed in Chapter 7 and derive the forms of the corresponding P^3 s by combining suitably the multiparameter hypotheses of 'no mixture'. The union-intersection tests derived with known p for two examples (one of the examples has already been considered in Chapter 6) are seen to serve as the appropriate P^3 tests when p also is unknown. For the last example we note that the appropriate P^3 coincides with the multivariate population skewness as derived by Mardia (1974) explicitly for multivariate normal mixtures and the test also coincides with the multivariate skewness test. In Section 9.6, the test is applied to a real-life example considered by Bacon-Shone and Fung (1987) in the context of detecting outliers. Part of this Chapter is based on Pal and SenGupta (2000b).

9.2. The LMMPU test

Let $g(x, \theta)$ be the density of a random observable X (may be vector-valued) whose distribution depends on a vector parameter $\theta = (\theta_1, \theta_2, \dots, \theta_k) \in \Theta \subset \mathcal{R}^k$. Consider the problem of testing the simple null hypothesis $H_0 : \theta = \theta_0$ against the alternative $H_1 : \theta \neq \theta_0$ where θ_0 is an interior point of Θ . Let $\phi = I_\omega$ be a test function for the above problem with critical region ω . Assume that the power function of ϕ , $\beta_\phi(\theta)$, is of class C^2 .

Definition 9.2.1. (SG-V) The level α test ϕ is (strictly) locally unbiased (LU) if $\beta_\phi(\theta) (>) \geq \alpha$ for all θ in some neighbourhood of θ_0 .

It can be noted that a test ϕ with $\beta_\phi(\theta)$ in C^2 is strictly LU of level α iff (1) $\beta_\phi(\theta_0) = \alpha$, (2) $\dot{\beta}_\phi(\theta_0) = 0$ and (3) $\ddot{\beta}_\phi(\theta_0)$ is positive definite. A set of necessary conditions for ϕ to be LU is (1), (2) and (3') $\ddot{\beta}_\phi(\theta_0)$ is nonnegative definite.

Definition 9.2.2. (SG-V) The level α test ϕ is LMMPU if it is LU of level α and for any other level α test ψ , there exists an $r_0 > 0$ (depending on ψ) such that

$$\int_{S_r} \beta_\phi(\theta) d\theta > \int_{S_r} \beta_\psi(\theta) d\theta, \quad r < r_0,$$

where $S_r = \{\theta : |\theta - \theta_0| < r\}$. Note also that a test ϕ is regular LMMPU with level α iff $\beta_\phi(\theta)$ satisfies (1) - (3) and (4) $\text{tr } \ddot{\beta}_\phi(\theta_0) \geq \text{tr } \ddot{\beta}_\psi(\theta_0)$ for any other level α test ψ . Replacing (3) by the weaker requirement (3') gives a second order LMMPU test.

Theorem 9.2.1. (SG-V, Theorem 2) Let $g(x, \theta)$, $\theta \in \Theta \subset \mathcal{R}^k$, be a k -parameter family of densities. Let $H_0 : \theta = \theta_0$ be a null hypothesis. Assume that the integral and derivative can be interchanged in β . Consider any Borel set ω of the form

$$\omega : \sum_{i=1}^k \ddot{g}_{ii}(x, \theta_0) \geq c g(x, \theta_0) + \sum_{i=1}^k c_i \dot{g}_i(x, \theta_0) \quad (9.2.1)$$

where the constants $c, c_1, \dots, c_k$ satisfy the conditions (1) $\int_\omega g(x, \theta_0) dx = \alpha$ and (2) $\int_\omega \dot{g}_i(x, \theta_0) dx = 0$ ($1 \leq i \leq k$). If ω is essentially the only set with the property (3) $\ddot{\beta}(\theta_0) = (\int_\omega \ddot{g}_{ij}(x, \theta_0)) \in \mathcal{R}^{k \times k}$ is positive definite, then ω is a regular LMMPU level α critical region. If $\ddot{\beta}(\theta_0)$ is only positive semidefinite, one has obtained a second order LMMPU level α critical region.

It may be observed that if $l(x, \theta) = \log g(x, \theta)$, the critical region (9.2.1) can be equivalently written as

$$\omega : \sum_{i=1}^k [\ddot{l}_{ii}(x, \theta_0) + \dot{l}_i^2(x, \theta_0)] \geq c + \sum_{i=1}^k c_i \dot{l}_i(x, \theta_0) \quad (9.2.2)$$

where $c, c_1, \dots, c_k \in \mathcal{R}$ are constants.

9.3. LMMPU tests in mixtures of univariate one-parameter distributions

We now apply the results of Section 9.2 to test the hypothesis of no mixture when n i.i.d. observations $(X_1, X_2, \dots, X_n)$ are drawn from a $(k+1)$ -component mixture of general one-parameter distributions. We assume

that each component of the mixture belongs to the same family of distributions given by the density f (w.r.t. some σ -finite measure). The joint density of the observations is then

$$g(x, \theta) = \prod_{j=1}^n [p_1 f(x_j, \theta_1) + p_2 f(x_j, \theta_2) + \dots + p_k f(x_j, \theta_k) + p_0 f(x_j, \theta_0)] \quad (9.3.1)$$

where $\theta = (\theta_1, \theta_2, \dots, \theta_k)$; $p_0 = 1 - p_1 - p_2 - \dots - p_k$; p_r 's are all known and θ_0 is also known. It is further assumed that the parameters are such that the mixture distribution is identifiable. The hypothesis of 'no mixture' can be expressed as a multiparameter hypothesis $H_0 : \theta = (\theta_0, \theta_0, \dots, \theta_0)' = \theta_0 \mathbf{1}$ which is to be tested against the alternative $H_1 : \theta \neq \theta_0 \mathbf{1}$ where $\mathbf{1}$ is the vector of all 1's. If $l(x, \theta) = \log g(x, \theta)$ be the log likelihood, then it follows that

$$\dot{l}_i(x, \theta_0) = p_i \sum_{j=1}^n \dot{q}(x_j, \theta_0) ,$$

$$\ddot{l}_{ii}(x_j, \theta_0) = p_i \sum_{j=1}^n \ddot{q}(x_j, \theta_0) + p_i(1 - p_i) \sum_{j=1}^n \dot{q}^2(x_j, \theta_0) ; i = 1, 2, \dots, k$$

where $q = \log f$ and $\dot{q}, \ddot{q}$ denote the first and second derivatives, respectively, of q w.r.t. θ at $\theta = \theta_0$. Writing $Z_1 = \sum_{j=1}^n \dot{q}(x_j, \theta_0)$, $Z_2 = \sum_{j=1}^n \dot{q}^2(x_j, \theta_0)$ and $Z_3 =$

$\sum_{j=1}^n \ddot{q}(x_j, \theta_0)$; $A_1 = \sum_{i=1}^k p_i^2$, $A_2 = \sum_{i=1}^k p_i(1 - p_i)$, $A_3 = \sum_{i=1}^k p_i$ and using (9.2.2), one can have the following

Theorem 9.3.1. For testing $H_0 : \theta = \theta_0 \mathbf{1}$ against $H_1 : \theta \neq \theta_0 \mathbf{1}$ in the model (9.3.1), the critical region given by

$$\omega = \left\{ (Z_1, Z_2, Z_3) : A_1 Z_1^2 + A_2 Z_2 + A_3 Z_3 \geq c + K Z_1 \right\} , \quad (9.3.2)$$

where K and c satisfy (i) the level α condition (ii) $\dot{\beta}_i(\theta_0 \mathbf{1}) = 0$ and (iii) $\ddot{\beta}(\theta_0 \mathbf{1})$ is positive definite (positive semidefinite), represents that of a regular (second order) LMMPU test of level α .

Remark 9.3.1. It is interesting to note that although there are k unknown constants $c_1, c_2, \dots, c_k$ in the second term of the right side of (9.2.2), in this case they get amalgamated to yield a single constant K . So one has

to determine the values of only two constants, *viz.*, K and c in order to obtain the LMMPU test. Thus some simplification is achieved. Also in some situations it turns out that $K = 0$ (we shall present one such example shortly). The only unknown is then c which can be determined from the level α condition. After having obtained c one then has to verify whether $\ddot{\beta}(\theta_0 \mathbf{1})$ is positive definite or positive semidefinite. Analytical verification is usually intractable. However, in some specific examples, this can be done numerically using a suitable computer package.

Illustration. As an illustration of Theorem 9.3.1, consider the mixture distribution in which $f(x, \theta_i)$ represents the density of a normal variable with mean θ_i and variance unity. Take $\theta_0 = 0$. Then $\dot{q} = x$, $\ddot{q} = -1$. It also follows that $Z_1 = \sum_{j=1}^n x_j = n\bar{x}$, $Z_2 = \sum_{j=1}^n x_j^2 = n(s^2 + \bar{x}^2)$ and $Z_3 = -n$. Write $u = \sqrt{n}\bar{x}$ and $v = ns^2$. Under H_0 , $u \sim N(0, 1)$, $v \sim \chi_{n-1}^2$ and u, v are independent with densities, say, $\tau(u)$ and $\xi(v)$ respectively. (9.3.2) is then equivalent to

$$\omega = \{(u, v) : Av + Bu^2 - Ku - d \geq 0\} \quad (9.3.3)$$

where A and B are known but K and d are unknown which are to be determined from the conditions of Theorem 9.3.1.

Result 9.3.1. In the critical region (9.3.3), $K = 0$ and hence (9.3.3) reduces to

$$\omega = \{(u, v) : Av + Bu^2 - d \geq 0\}. \quad (9.3.4)$$

Proof. We shall exploit the reflection principle as enunciated in Ferguson (1967, pp. 239-240) to accomplish the proof. Note that conditions (i) and (ii) of Theorem 9.3.1 are, respectively, equivalent to

$$\int_{\omega} \tau(u) \xi(v) du dv = \alpha \quad (9.3.5)$$

and

$$\int_{\omega} u \tau(u) \xi(v) du dv = 0. \quad (9.3.6)$$

One can write $\omega = \omega_1 \cup \omega_2$ where

$$\omega_1 = \{(u, v) : Av + Bu^2 - d \geq Ku, u \geq 0\}$$

and

$$\omega_2 = \{(u, v) : Av + Bu^2 - d \geq Ku, u \leq 0\}.$$

Let ω'_2 be the image of ω_2 under reflection through the $\sqrt{v}$ -axis, so that

$$\omega'_2 = \{(u, v) : Av + Bu^2 - d \geq -Ku, u \geq 0\}.$$

If $K \leq 0$, then clearly $\omega'_2 \subset \omega_1$. Hence from (9.3.6)

$$\begin{aligned} 0 = \int_{\omega} u \tau \xi \, du \, dv &= \int_{\omega_1 \cup \omega_2} u \tau \xi \, du \, dv \\ &= \int_{\omega_2} u \tau \xi \, du \, dv + \int_{\omega'_2} u \tau \xi \, du \, dv + \int_{\omega_1 \setminus \omega'_2} u \tau \xi \, du \, dv \\ &= \int_{\omega_1 \setminus \omega'_2} u \tau \xi \, du \, dv. \end{aligned}$$

But $\omega_1 \setminus \omega'_2 = \{(u, v) : d + Ku \leq Av + Bu^2 \leq d - Ku, u \geq 0\}$. Also the integrand in the last integral over the set $\omega_1 \setminus \omega'_2$ is nonnegative. Hence this region of integration must collapse to a null set, *i.e.*, K must be zero. The same argument is valid if $K \geq 0$. Hence the proof. $\square$

The only unknown constant in (9.3.4) is now d . This should be determined from (9.3.5). Some numerical procedure may be adopted for this purpose. Analytical verification of the definiteness (positive or nonnegative) of $\tilde{\beta}(\mathbf{0})$, even for this example, is extremely difficult. One can, however, use some standard package for the computation of each element of this matrix numerically and then verify its definiteness.

9.4. LMMPU tests in multivariate mixtures

Let us now consider the problem of testing the hypothesis of no contamination in a two-component mixture of multivariate distributions each depending on a k -component vector parameter of which one is, of course, known. The joint density of a random sample $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$ for such a model is given by

$$g(\mathbf{x}, \boldsymbol{\theta}) = \prod_{j=1}^n \{p f(\mathbf{x}_j, \boldsymbol{\theta}) + (1 - p) f(\mathbf{x}_j, \boldsymbol{\theta}_0)\} \quad (9.4.1)$$

where each X_j is the associated random vector of order, say $r \times 1$, and θ is a vector of unknown parameters belonging to $\Theta \subset \mathcal{R}^k$; p ($0 < p < 1$) is known and θ_0 is assumed to be a known point of Θ . To test $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$ the general form of the LMMPU critical region can be derived from (9.2.2). We, however, consider below the case, as an example, when $f(x, \theta)$ represents the density of a k -variate normal distribution.

9.4.1. Example Let $f(x, \theta)$ be the density of a k -variate normal distribution with mean vector θ and covariance matrix I_k . We then have $r = k$. Take $\theta_0 = 0$. With the same notations as in the earlier Section, one can obtain the following after some routine algebra:

$$\dot{l}_i = np\bar{x}_i ; \quad \ddot{l}_{ii} = npq(s_{ii} + \bar{x}_i^2) - np,$$

where x_{ij} is the i -th component of x_j , $\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ij}$, $s_{ii} = \frac{1}{n} \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2$ and $q = 1 - p$. Also define $s_{im} = \frac{1}{n} \sum_{j=1}^n (x_{ij} - \bar{x}_i)(x_{mj} - \bar{x}_m)$. Then $\ddot{l}_{im} = np(1-p)(s_{im} + \bar{x}_i\bar{x}_m)$. Writing $A = p(1-p)$, $B = p(1-p) + np^2$, $u_i = \sqrt{n}\bar{x}_i$, $v_i = ns_{ii}$ and making some simplifications, it follows that (9.2.2) takes the form

$$\omega = \left\{ (u, v) : A \sum_{i=1}^k v_i + B \sum_{i=1}^k u_i^2 \geq c + \sum_{i=1}^k D_i u_i \right\}. \quad (9.4.2)$$

where $u = (u_1, u_2, \dots, u_k)$ and $v = (v_1, v_2, \dots, v_k)$.

Let $\tau(u)$ and $\xi(v)$ now denote the densities of u and v respectively under H_0 . We shall show in the following theorem that condition (2) of Theorem 9.2.1 is equivalent to $D_i = 0$ for all i .

Theorem 9.4.1. Condition (2) of Theorem 9.2.1 is satisfied iff $D_i = 0$ for all $i = 1, 2, \dots, k$.

Proof. For any fixed i if $D_i = 0$, then ω is a symmetric region w.r.t. u_i . Hence

$$\dot{\beta}_i(0) = \int_{\omega} u_i \tau(u) \xi(v) du dv = 0$$

since the integrand is an odd function of u_i .

Conversely, suppose $\dot{\beta}_i(0) = 0$ for any given i . Note that ω can be written as $\omega = \omega_1 \cup \omega_2$, where

$$\omega_1 = \left\{ (u, v) : A \sum_i v_i + B \sum_i u_i^2 - c \geq \sum_i D_i u_i, u_i \geq 0 \right\},$$

$$\omega_2 = \left\{ (u, v) : A \sum_i v_i + B \sum_i u_i^2 - c \geq \sum_i D_i u_i, \quad u_i \leq 0 \right\}.$$

Let ω'_2 be the image of ω_2 under the reflection along $u_i = 0$. That is

$$\omega'_2 = \left\{ (u, v) : A \sum_i v_i + B \sum_i u_i^2 - c \geq -D_i u_i + \sum_{i' \neq i} D_{i'} u_{i'}, \quad u_i \geq 0 \right\}.$$

If $D_i \leq 0$, then $\omega'_2 \subset \omega_1$. Proceeding exactly as in Section 9.3, it follows that

$$0 = \int_{\omega} u_i \tau(u) \xi(v) du dv = \int_{\omega_1 \setminus \omega'_2} u_i \tau(u) \xi(v) du dv,$$

which must lead to $D_i = 0$. Similarly for the other case and hence the proof. $\square$

In the light of Theorem 9.4.1, the critical region ω takes the form

$$\omega = \left\{ (u, v) : A \sum_{i=1}^k v_i + B \sum_{i=1}^k u_i^2 \geq c \right\}. \quad (9.4.3)$$

Under H_0 , $V = \sum_i v_i \sim \chi_{k(n-1)}^2$, $U^2 = \sum_i u_i^2 \sim \chi_k^2$ and U and V are independently distributed. The only constant c is to be determined from the level α condition using a computer package. Analytical verification of positive definiteness of $\ddot{\beta}(0)$ is again intractable and is to be done numerically.

Remark 9.4.1. As expected, for $p = 1$, the critical region ω of (9.4.3) coincides with the LMMPU critical region of SG-V derived for testing the hypothesis that the mean vector of a k -variate normal distribution with covariance matrix I_k is the null vector.

9.5. P^3 tests for multiparameter hypotheses in multivariate mixtures

In this Section we consider multivariate analogs of the univariate mixture models (i), (ii) and (iii) of Section 7.7 of Chapter 7. The corresponding densities are

$$(i'') \quad g(x|p, \mu_1, \mu_2) = pN_k(x|\mu_1, I) + (1-p)N_k(x|\mu_2, I)$$

$$(ii'') \quad g(x|p, \mu, \Sigma) = pN_k(x|\mu, \Sigma) + (1-p)N_k(x|0, \Sigma)$$

(iii'') $g(x|p, \mu_1, \mu_2, \Sigma) = pN_k(x|\mu_1, \Sigma) + (1-p)N_k(x|\mu_2, \Sigma)$ (Ganesalingam and McLachlan, 1981).

For each of the above models our concern is to derive appropriate test for the hypothesis of 'no mixture' which is clearly a multiparameter hypothesis. To achieve this we combine the optimal tests obtained for the corresponding univariate models by the $\cup\cap$ principle of Roy (1953) (see Chapter 6). We first consider the case when $p \in (0, 1)$ is known and then move on to the case of unknown p .

Theorem 9.5.1. (a) An $\cup\cap$ test for the hypothesis of 'no mixture' for model (i'') consists in rejecting the hypothesis for large values of $\lambda_M(A)$ where $A = \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})'$ and $\lambda_M(A)$ is the largest characteristic root of A .

(b) An $\cup\cap$ test for the hypothesis of 'no mixture' for model (ii'') consists in rejecting the hypothesis for large values of $\bar{X}'A^{-1}\bar{X}$.

(c) An $\cup\cap$ test for the hypothesis of 'no mixture' for model (iii'') consists in rejecting the hypothesis for large values of $\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n [(X_i - \bar{X})'S^{-1}(X_j - \bar{X})]^2$ where $S = A/n$.

Proof. First consider (a). Taking L to be the collection of all k -vectors l of unit norm and appealing to the corresponding univariate model, it follows that

$$H_0 : \text{'no mixture'} \equiv \bigcap_{l \in L} \left\{ H_{0l} : p(1-p) [l'(\mu_1 - \mu_2)]^2 = 0 \right\}.$$

Note that

$$\sup_{l \in L} [l'(\mu_1 - \mu_2)]^2 = (\mu_1 - \mu_2)'(\mu_1 - \mu_2).$$

Again H_{0l} is rejected if $\sum_{i=1}^n (Y_i - \bar{Y})^2 > c_l$ where $Y_i = l'X_i$, $i = 1, 2, \dots, n$.

By the $\cup - \cap$ principle of Roy (1953), H_0 is therefore rejected if $T'_1 = \sup_{l \in L} l'Al > c$ i.e., if $\lambda_M(A) > c$. The cut-off point c can be found from Sugiyama (1972) or from Table 51 of Pearson and Hartley (1972).

Next consider (b). Proceeding exactly as in (a), the corresponding $\cup\cap$ test is seen to reject H_0 whenever $T'_2 = \bar{X}'A^{-1}\bar{X} > c$. The cut-off point c

can be calculated from the table of F -distributions, since under H_0 , the distribution of $\bar{X}'A^{-1}\bar{X}$ is that of a Hotelling's T^2 statistic which has relation with an F -statistic with appropriate degrees of freedom.

Finally for (c), it follows by the same reasoning as in case (a), that the $\cup\cap$ test consists in rejecting H_0 for large values of

$$T'_3 = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n [(X_i - \bar{X})'S^{-1}(X_j - \bar{X})]^3. \quad \square$$

Interestingly, we note that the P^3 s for each of the above three cases reduces to the following:

$$\eta'_1 \equiv p(1-p)(\mu_1 - \mu_2)'(\mu_1 - \mu_2);$$

$$\eta'_2 \equiv \mu'[\Sigma + p(1-p)\mu\mu']^{-1}\mu;$$

and

$$\eta'_3 \equiv \frac{a(p)\Delta^6}{(1+p(1-p)\Delta^2)^3}.$$

Remark 9.5.1. Recall that the above expression for η'_3 coincides with the population skewness in (2.14) of Mardia (1974) derived explicitly for normal mixtures and the test statistic T'_3 is the sample measure of multivariate skewness used for testing normality (Mardia, 1974) and also for the purpose of discrimination (McLachlan, 1992, pp. 172-173). Exact cut-off points for the particular case of $k = 2$ can be found from Mardia (1974). The asymptotic distribution of the test statistic of that paper may also be used to obtain the asymptotic cut-off points.

Next suppose p is unknown. In this case we have to define the P^3 s appropriately and construct tests based on their unbiased and/or consistent estimators. We see that the η'_i s of the previous case can be taken as the P^3 s and also the corresponding test statistics T'_i s, yielding the $\cup\cap$ tests, are equivalent to consistent estimators in the respective models. These tests have the property that for each given $p \in (0, 1)$ they are $\cup\cap$ tests obtained from the corresponding univariate models. We then have the following

Theorem 9.5.2. The $\cup\cap$ tests of Theorem 9.5.1 are the appropriate P^3 -tests corresponding to the P^3 s η'_i s as defined in the proof of the theorem.

9.6. Illustrative example

Bacon-Shone and Fung (1987) analyzed the milk transportation-cost data of Johnson and Wichern (1982, p. 282) for the detection of multivariate outliers using some graphical approach under the assumption of multivariate normality. They considered costs corresponding only to the gasoline trucks and discovered that out of 36 trivariate sample observations two, *viz.*, the 9th and the 21st, are outliers (see also Johnson and Wichern, 1996, pp. 276-277). It is therefore quite reasonable to assume a mixture model of the form (iii'') for these data. Further it is also reasonable to assume that p , the mixing proportion, is quite different from $1/2$. We now apply the multivariate skewness test on all these 36 observations and note that the calculated value of T'_3 comes out as 6.3185. The value of the approximating χ^2 statistic, as given in Mardia (1974), is 37.911 which is highly significant as compared to $\chi^2_{10,05} = 18.31$.

Next we delete those two outlying observations and see for the remaining 34 observations that $T'_3 = 1.1255$. Also $\chi^2 = 6.378$ which is much less than the 5% critical value. Thus the skewness test, while successfully identifying the population distribution as a mixture of two multivariate normals, strongly supports the findings of Bacon-Shone and Fung (1987).

Chapter 10

An Overview – Directions for Further Research

The present thesis is an attempt to explore a few possible approaches to exact and asymptotically optimal tests for no contamination in mixtures of distributions. While the exact optimality criterion is the maximization of local power based on Neyman- Pearson framework, we use the approaches of Neyman (1959), Davies (1977, 1987) and Bahadur (1960, 1967) for the asymptotic tests. In the multiparameter situations, the directional derivative test of John (1971) may also be tried.

There is another approach of optimal testing with the criterion of ‘beta optimality’ introduced by Davies (1969) and further considered by, *e.g.*, King (1987) and Bhatti and King (1990). A level α test ψ is said to be *beta optimal* if among all level α tests with monotone power function, ψ attains a preassigned power, say β , *most quickly*. King (1987) has shown that if a point optimal test (*i.e.*, a test which is most powerful against a given simple alternative) has a monotone power function then this test is also beta optimal. Application of this approach to mixture models even for specific examples like normal mixtures to obtain exact analytic tests, however, seems to be formidable. One, of course, may try for some approximations to obtain workable tests in this regard. The author plans to work on this topic in future.

In Chapter 3, although we have considered mixtures of normals, the technique can be applied quite generally to any distribution showing non-regularities of the types discussed in this Chapter. Recently, in the context of a different problem, Yamamoto and Shinozaki (2000) considered a ‘mirrored mixture’ of the form

$$f_m(x) = \frac{1}{2}g(x + m) + \frac{1}{2}g(x - m)$$

where g is a density symmetric about 0. The nature of reparametrization will, of course, depend on the form of g .

The approach of Davies leads to conservative tests. As Titterington *et al.* (1985) have remarked, "In principle, the method can be extended to deal with problems involving more than one nuisance parameter, such as more complicated mixture models", this approach has further scope to apply in non-regular situations.

The test procedures used in Chapter 4 are the locally best tests or their different modifications. One reason for considering such tests is that they are easy to obtain and, as is seen, enjoy nice properties, *viz.*, monotonicity of the power function and consistency. Also it should be borne in mind that it is, in general, difficult to detect small departures from the null hypotheses while large departures can be detected quite easily by any reasonable test. The LRTs in all these situations are very computation-intensive and difficult to apply since they cannot be written in any closed form. Note that no nontrivial sufficient statistic exists for our mixture family and hence this is expected. The exact distribution of the LRT statistic is intractable. Further, the standard result yielding the χ^2 as the asymptotic null distribution of the LRT statistic is not valid for our case. This is so due to the non-regular nature of our problem where the parameter space is no longer open and where the parameter lies on the boundary under H_0 . For this reason the LRT approach is not to be pursued and any numerical comparison, which can be possibly explored at most via simulations, seems unappealing and uninformative. However, it may be worthwhile to note that the LMP test can be viewed as a first order approximation of the LRT.

The rose diagram of the solar wind data in Figure 11.3.2 of Chapter 11 looks slightly bipolar. We plan to analyze these data using a suitable bipolar distribution in future.

The problem of detection of outliers in directional data analysis is currently drawing attentions of many researchers. The review paper by Jupp and Mardia (1989) contains, among others, an extensive references on this interesting area. Guttorp and Lockhart (1988) provide a Bayesian solution of the problem of detecting the location of a downed aircraft from distress signals transmitted by it and received by different search-and-rescue stations. Outliers in the data may occur due to irregularities in the readings caused by distorting objects near the receptor site. One can, therefore, ei-

ther detect and reject the outliers or can assume a mixture model in order to incorporate them for the analysis. Thus if one intends to ‘accommodate’ outliers in the model, it is a common practice to assume a mixture distribution for it. Otherwise one may assume a non-mixture model like CN (Collett, 1980) for outlier detection. The SWS distribution considered in Section 4.3.2 of Chapter 4 may also be used for this purpose and this may be an interesting problem for further research.

Let us now consider Chapter 5. For the general problem for no mixture, one would like to characterize this hypothesis by translating it to a unique (except for the nuisance parameters) hypothesis on the parameters. However, to avoid the problem of non-identifiability, it might be necessary to make some assumptions, leading to special cases. One special case is that of the ‘strongly identifiable’ mixture (Ghosh and Sen, 1985) where this is achieved through restrictions on the form of the component densities, *e.g.*, requiring that they be of different families or of different types, $g_p(x) = pf_1(x, \theta) + qf_2(x, \eta)$. In most applications, however, as Titterton *et al.* (1985, p. 6) have pointed out, we “would mainly be considering mixtures of component densities all belonging to the same parametric family” and all of the same type. Then, $g_p(x) = pf(x, \theta_1) + qf(x, \theta_2)$, *e.g.*, mixture of normals, mixture of exponentials etc. In this case, the hypothesis H_0 : ‘no mixture’, is equivalent to either $H_0 : p = 1$ or $H_0 : \theta_1 = \theta_2$ or $H_0 : p = 1, \theta_1 = \theta_2$, simultaneously. Here one would achieve the aforesaid desired characterization by assuming, as another special case, that either $p \neq 1$ or $H_0 : \theta_1 \neq \theta_2$. In many cases, practical considerations may suggest that $p \neq 1$. Further, it may be natural for many models to have θ_1 known, say $\theta_1 = \theta_{10}$; *e.g.*, in Hartigan’s problem, $\theta_1 \equiv \mu, \theta_{10} = 0$; in our problem, $\theta_1 \equiv \rho, \theta_{10} = 0$; etc. Then the assumption of p known leads to the LMP test for H_0 : ‘no mixture’ in g_p , to coincide with that for $H_0 : \theta_1 = \theta_{10}$ in $f(x, \theta)$. Thus, often the test statistic is quite simple and exact and elegant results regarding its properties may be obtained.

If p is also a nuisance parameter, one can attempt to construct asymptotically LMP tests, *e.g.*, Neyman’s $C(\alpha)$ test, likelihood ratio derivative test, etc. Further when either or both ρ and q may assume zero values, the test may be based on an appropriate Pivotal Parametric Product, as discussed in Chapter 7. Details may be pursued in the lines of Rachev and SenGupta (1992) and Pal and SenGupta (2000a). Finally, even if p is to

be tested also, one can explore the multiparameter LMMPU tests of Sen-Gupta and Vermeire (1986), *e.g.*, to test, for our problem, $H_0 : p = p_0$ and $\rho = \rho_0$, $p_0 \in (0, 1)$, $\rho_0 \in (-1/(k-1), 1)$, simultaneously. Our approach, in the given framework, is thus applicable quite generally and the above observations are interesting topics for further research.

The P^3 approach considered in Chapter 7 is quite simple and leads to tests which are optimal in a number of senses as discussed in that Chapter. The consistency of the test statistic for the appropriate P^3 ensures that the test is consistent — which is a desirable criterion in large samples. This approach in mixtures involving more than two components, however, seems not to be so easy and may be explored for future research.

The test proposed in Section 9.3 of Chapter 9 is one member of the class of locally Bahadur optimal tests and there may be many such tests having the desired property. In order to ensure local Bahadur optimality of the proposed test, one, however, has to verify assumption (A3) separately for each member of the exponential family. For the multiparameter exponential family of distributions, analogous results may possibly be established and these may be interesting topics for further research.

Chapter 11

Computational Support and Results

We present in this Chapter some selected computer programs and all the relevant tables and diagrams. The diagrams have been prepared using the statistical package DDSTAP (SenGupta, 1998b). The computer programs have been written in FORTRAN.

11.1. Some selected computer programs

```
c *** Program 11.1.1
c *** This program calculates the simulated power of the
c *** LMPI test as a function of p for various combinations
c *** of m, a and rho
      dimension prop(10),jj1(10),stat(10000),power5(10),
1 theta(50),jj2(10),power1(10)
      integer x,y
      external statcomp,rnun,rnsta
      write(*,*)'Give sample size'
      read(*,*) m
      write(*,*)'Give a'
      read(*,*) a
      write(*,*)'Give rho'
      read(*,*) rho
      write(*,*)'Give 5% critical value corresponding to the
1 above combination of m, a and rho'
      read(*,*) cut05
      write(*,*)'Give 1% critical value corresponding to the
1 above combination of m, a and rho'
      read(*,*) cut01
      write(*,*)'Give initial value of the random number seed'
```

```

    read(*,*) iseed
    write(*,*) 'How many times you want to generate your
1 statistic ?'
    read(*,*) nnn
    do 700 i=1,10
    prop(i)=float(i)/10.0
    do 490 j=1,nnn
    do 480 k=1,m
    call rnun(1,r)
    iseed=iseed+1
    if(r.lt.prop(i)) go to 470
    call rnun(1,angle)
    theta(k)=angle*6.28318
    go to 120
470  call rnsta(1,a,0.0,angle1)
    angle1=angle1*((-alog(rho))**(1./a))
    if(angle1.lt.0.0) go to 765
    angle1=amod(angle1,6.28318)
    theta(k)=angle1
    go to 120
765  angle1=amod(angle1,6.28318)+6.28318
    theta(k)=angle1
120  iseed=iseed+1
480  continue
    call statcomp(m,a,rho,theta,stat(j))
    iseed=iseed+1
490  continue
    jj1(i)=0
    do 500 x=1,nnn
    if(stat(x).le.cut05) go to 500
    jj1(i)=jj1(i)+1
500  continue
    jj2(i)=0
    do 600 y=1,nnn
    if(stat(y).le.cut01) go to 600
    jj2(i)=jj2(i)+1

```



```

600  continue
      power5(i)=float(jj1(i))/float(nnn)
      power1(i)=float(jj2(i))/float(nnn)
700  continue
      write(*,*)' m=',m,' rho=',rho,' a=',' cutoff point
1  at 5%=',cut05,' cutoff point at 1%=',cut01
      write(*,*)' p', ' power at 5%', ' power at 1%'
      do 900 ll=1,10
      write(*,45) prop(ll),power5(ll),power1(ll)
45  format(1x,f3.1,f8.4,5x,f8.4)
900  continue
      stop
      end
c *** This subroutine calculates the value of the
c *** LMPI test statistic
      subroutine statcomp(mm,a,r,thet,st)
      dimension thet(mm),cs(50),sn(50),rnsq(50)
      do 100 n=1,30
      cs(n)=0.
      sn(n)=0.
      rnsq(n)=0.
      do 200 j=1,mm
      cs(n)=cs(n)+cos(n*thet(j))
      sn(n)=sn(n)+sin(n*thet(j))
200  continue
      rnsq(n)=cs(n)*cs(n)+sn(n)*sn(n)
      rnsq(n)=((r*r)**(n**a))*rnsq(n)
100  continue
      st=0.
      do 300 n=1,30
      st=st+rnsq(n)
300  continue
      return
      end

```

```

c  ** Program 11.1.2
c  ** This program calculates relevant test statistics
c  ** for wrapped stable model for ant and solar wind data
    implicit real*8(a-h,o-z)
    dimension angle(40),frq(40),rnsq1(40),rnsq2(40),rnsq(40)
    write(*,*)' What is the value of a ?'
    read(*,*)a
    write(*,*)' What is rho ?'
    read(*,*)rho
    write(*,*)'How many classes in the freq. distn. ?'
    read(*,*)nclass
    do 10 i=1,nclass
    read(5,*)angle(i),frq(i)
    angle(i)=angle(i)*3.14159/180.0
10  continue
    totfr=0.
    do 20 i=1,nclass
20  totfr=totfr+frq(i)
    write(4,*)' Total fr. is ',totfr
    write(4,*)
    write(4,*)' Following are the Rnsq. values '
    do 30 i=1,30
    rnsq1(i)=0.
    rnsq2(i)=0.
    do 40 j=1,nclass
    rnsq1(i)=rnsq1(i)+frq(j)*dcos(i*angle(j))
40  rnsq2(i)=rnsq2(i)+frq(j)*dsin(i*angle(j))
    cbar=rnsq1(i)/totfr
    sbar=rnsq2(i)/totfr
    rnsq(i)=cbar*cbar+sbar*sbar
    write(4,*) rnsq(i)
30  continue
    statlm=0.0
    do 50 i=1,30
50  statlm=statlm+((rho*rho)**(i**a))*rnsq(i)*2.0*totfr
    ralest=rnsq(1)*2.0*totfr

```

```

write(4,*)'LMPI statistic has the value  '
write(4,*)statlm
write(4,*)
write(4,*)  'Rayleighs statistic has the value  '
write(4,*) ralest
stop
end

```

c**** Program 11.1.3

c**** Calculation of significance probabilities

c**** of the LMPI and the Davies-motivated tests

```

implicit real*8(a-h,o-z)

```

```

dimension a(100),b(100)

```

```

write(*,*) ' Give the value of a'

```

```

read(*,*) aind

```

```

write(*,*) ' Give the value of rho'

```

```

read(*,*) rho

```

```

write(*,*) 'Give the value of the Statistic'

```

```

read(*,*) t

```

```

do 10 i=1,25

```

```

a(i)=1.

```

```

do 20 j=1,25

```

```

if(j.ne.i) go to 50

```

```

b(j)=1.

```

```

go to 60

```

```

50 b(j)=1./(1.-(rho*rho)**((j**aind)-(i**aind)))

```

```

60 a(i)=a(i)*b(j)

```

```

20 continue

```

```

10 continue

```

```

prob=0.0

```

```

do 70 i=1,25

```

```

prob=prob+a(i)*exp(-t/(2.*(rho*rho)**(i**aind)))

```

```

70 continue

```

```

write(*,23) prob

```

```

23 format(' The reqd. significance prob. is ',e20.8)

```

```
stop
end
```

```
c *** Program 11.1.4
c *** This program calculates the statistical curvature
c *** for various combinations of q and k
      dimension gam(11,20),g(20)
      do 10 i=5,95,10
      do 20 k=2,10,2
      gam(i,k)= 8./float(k)+
2 float(3*k*k-7*k+6)*(((100.-float(i))/float(i))**2)
3   * 2./float(k*(k-1))
20   continue
10   continue
      do 30 i=5,95,10
      write(*,40) (gam(i,k), k=2,10,2)
30   continue
40   format(5f10.2)
      do 70 k=2,10,2
      g(k)= 8./float(k)
70   continue
      write(*,40) (g(k),k=2,10,2)
      stop
      end
```


11.2. Tables

Table 11.2.1. Simulated cut-off points of the LMPI test based on T

m	a ρ	0.5	1.0	1.5	2.0
10	.25	(2.25,3.30)	(1.86,2.74)	(1.81,2.73)	(1.78,2.59)
	.50	(12.94,16.81)	(8.40,11.48)	(7.53,10.73)	(7.26,10.56)
	.75	(53.74,65.48)	(25.12,35.99)	(19.78,27.92)	(17.00,24.90)
	.90	(124.97,157.95)	(64.86,79.08)	(39.55,52.67)	(30.52,43.53)
15	.25	(3.36,4.67)	(2.84,4.17)	(2.69,4.05)	(2.61,4.00)
	.50	(19.33,25.85)	(12.96,17.84)	(11.74,17.64)	(11.35,16.26)
	.75	(81.41,96.91)	(37.46,50.37)	(28.29,40.89)	(26.78,39.99)
	.90	(189.15,225.82)	(97.02,119.74)	(58.53,81.79)	(46.15,60.80)
20	.25	(4.44,6.29)	(3.81,5.84)	(3.77,5.69)	(3.59,5.45)
	.50	(25.82,35.90)	(17.48,24.73)	(16.08,23.05)	(14.77,21.71)
	.75	(110.06,133.34)	(51.53,65.28)	(37.80,51.91)	(35.61,50.21)
	.90	(250.27,300.33)	(127.81,159.79)	(78.08,101.63)	(63.26,85.76)
30	.25	(6.54,9.42)	(5.66,8.35)	(5.29,8.09)	(5.27,7.67)
	.50	(40.17,52.45)	(24.66,36.26)	(23.44,35.44)	(21.13,30.38)
	.75	(163.61,199.44)	(78.37,105.04)	(62.08,94.85)	(53.25,82.42)
	.90	(376.02,442.52)	(196.16,240.38)	(119.42,156.11)	(94.53,131.99)

(The figures inside the brackets denote the cut-off points at 5% and 1% levels respectively.)

Table 11.2.2. Simulated power of the LMPI test at 5% level

α	ρ	p m	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
.5	.25	10	.051	.054	.055	.068	.077	.082	.106	.119	.138	.156
		20	.054	.061	.072	.098	.108	.145	.165	.200	.256	.300
	.50	10	.053	.063	.090	.145	.193	.251	.337	.413	.518	.605
		20	.062	.095	.168	.248	.356	.482	.635	.734	.828	.921
	.75	10	.072	.120	.237	.363	.502	.645	.783	.871	.942	.972
		20	.078	.214	.390	.638	.827	.913	.971	.993	1.000	1.000
	.90	10	.074	.215	.396	.592	.760	.877	.944	.986	.998	1.000
		20	.142	.400	.689	.898	.963	.993	.999	1.000	1.000	1.000
1.0	.25	10	.054	.057	.058	.060	.068	.093	.094	.112	.130	.144
		20	.056	.058	.073	.079	.099	.123	.133	.183	.236	.282
	.50	10	.055	.063	.077	.100	.146	.191	.251	.335	.419	.502
		20	.058	.067	.105	.169	.247	.359	.473	.604	.723	.858
	.75	10	.062	.088	.157	.262	.367	.522	.635	.783	.874	.949
		20	.065	.140	.255	.442	.644	.826	.911	.973	.994	1.000
	.90	10	.074	.134	.283	.449	.628	.779	.897	.958	.992	1.000
		20	.106	.295	.535	.766	.912	.978	.998	1.000	1.000	1.000
1.5	.25	10	.052	.053	.060	.067	.081	.084	.088	.108	.131	.149
		20	.053	.057	.068	.078	.089	.117	.141	.168	.221	.271
	.50	10	.055	.061	.082	.114	.150	.207	.272	.336	.440	.532
		20	.056	.070	.106	.158	.235	.364	.477	.609	.744	.838
	.75	10	.056	.074	.127	.214	.310	.425	.589	.734	.847	.946
		20	.073	.145	.259	.401	.599	.773	.904	.974	.993	1.000
	.90	10	.073	.115	.212	.368	.524	.702	.846	.929	.987	.999
		20	.091	.200	.423	.661	.866	.945	.989	.999	1.000	1.000
2.0	.25	10	.053	.056	.062	.071	.075	.080	.100	.126	.132	.157
		20	.054	.062	.071	.087	.109	.128	.157	.201	.235	.288
	.50	10	.054	.057	.078	.111	.158	.199	.269	.353	.443	.539
		20	.055	.081	.127	.169	.289	.429	.520	.663	.786	.881
	.75	10	.060	.098	.140	.205	.335	.463	.615	.762	.900	.981
		20	.063	.123	.223	.386	.574	.746	.891	.975	.997	1.000
	.90	10	.068	.109	.198	.339	.492	.652	.830	.934	.988	1.000
		20	.093	.206	.329	.584	.816	.939	.982	1.000	1.000	1.000

Table 11.2.3. Results of the application of optimal tests on Ant data

Test applied	p-value (significance probability)	Remark
LMPI	0.0	Highly significant
Rayleigh's R^2	0.0	Highly significant
Davies-motivated ($\epsilon = 0.1$)	0.0	Highly significant

Table 11.2.4. Goodness-of-fit tests for Jander's Ant data

Distribution fitted	U^2	Remark
CN	.4799	Not satisfactory
SWS ($a = .8$)	.0492	Satisfactory
MIX ($a = 1.59, p = .7$)	.039	More satisfactory

Table 11.2.5. Results of the application of optimal tests on solar wind data

Test applied	p-value (significance probability)	Remark
LMPI	0.78087×10^{-5}	Highly significant
Rayleigh's R^2	0.00098	Highly significant
Davies-motivated ($\epsilon = 0.1$)	0.0	Highly significant

Table 11.2.6. Goodness-of-fit tests for solar wind data

Distribution fitted	U^2	Remark
CN	.1248	Satisfactory
SWS ($a = .47$)	.0694	More satisfactory
MIX ($a = 2.0, p = .3$)	.0542	Much more satisfactory

Table 11.2.7. Values of ${}_1\gamma_{\rho=0}^2$ with $q = 0.05(0.10)1.00$
and $k = 2(2)10$

k q	2	4	6	8	10
0.05	1448.00	1566.33	1734.13	1831.79	1894.04
0.15	132.44	141.15	155.47	163.85	169.20
0.25	40.00	41.00	44.53	46.64	48.00
0.35	17.80	16.95	17.89	18.49	18.89
0.45	9.98	8.47	8.50	8.58	8.63
0.55	6.68	4.90	4.55	4.39	4.31
0.65	5.16	3.26	2.73	2.47	2.32
0.75	4.44	2.48	1.87	1.56	1.38
0.85	4.12	2.13	1.48	1.16	0.96
0.95	4.01	2.01	1.35	1.01	0.81
1.00	4.00	2.00	1.33	1.00	0.80

Table 11.2.8. Exact power values of the LMP test
 $(\alpha = 0.05, k = 3; n = 10 \text{ and } 25)^1$

p ρ	0.10	0.30	0.50	0.70	0.90
0.10	0.1210	0.1074	0.0923	0.0689	0.0593
	0.2011	0.1533	0.1257	0.0870	0.0577
0.30	0.3750	0.2832	0.1989	0.1361	0.0709
	0.6589	0.5117	0.3476	0.1976	0.0908
0.50	0.6354	0.4785	0.3354	0.2047	0.0945
	0.9298	0.8103	0.5870	0.3475	0.1284
0.70	0.8188	0.6753	0.4791	0.2856	0.1168
	0.9911	0.9384	0.7786	0.4912	0.1759
0.90	0.9121	0.7848	0.5944	0.3635	0.1424
	0.9991	0.9841	0.8827	0.6205	0.2214
0.99	0.9409	0.8260	0.6454	0.4051	0.1602
	0.9998	0.9907	0.9154	0.6531	0.2275

¹First and second sets of entries correspond to $n = 10$ and $n = 25$ respectively

Table 11.2.9. Power of T_2 -test with $\alpha = 0.05$.

$\rho \backslash p$	.1	.2	.3	.4	.5	.6	.7	.8	.9
.1	.2426	.2276	.2281	.2158	.2012	.1889	.1794	.1697	.1645
.2	.3490	.3202	.2941	.2857	.2475	.2399	.2100	.1946	.1663
.3	.4251	.3986	.3726	.3304	.3186	.2774	.2400	.2173	.1872
.4	.5109	.4851	.4456	.3864	.3505	.3186	.2791	.2275	.1968
.5	.5893	.5479	.4943	.4620	.3997	.3428	.2951	.2526	.2046
.6	.6407	.6121	.5565	.5104	.4579	.4090	.3381	.2798	.2084
.7	.6885	.6392	.6059	.5359	.4835	.4318	.3804	.2878	.2203
.8	.7314	.6879	.6487	.5802	.5283	.4455	.4050	.3148	.2487
.9	.7603	.7265	.6816	.6213	.5593	.4867	.4194	.3474	.2528

Table 11.2.10. Power of T_3 -test with $\alpha = 0.05$.

$\rho \backslash p$	.1	.2	.3	.4	.5	.6	.7	.8	.9
.1	.1394	.1157	.1074	.1015	.0956	.0835	.0868	.0620	.0620
.2	.2414	.2227	.1924	.1621	.1452	.1188	.0998	.0837	.0669
.3	.3845	.3441	.2837	.2502	.2104	.1733	.1424	.1134	.0789
.4	.5482	.4825	.4201	.3538	.2802	.2312	.1709	.1319	.0829
.5	.6996	.6388	.5506	.4390	.3663	.2914	.2108	.1560	.0999
.6	.8275	.7415	.6608	.5485	.4529	.3500	.2646	.1747	.1133
.7	.9108	.8317	.7424	.6473	.5255	.4257	.3055	.2025	.1228
.8	.9621	.8984	.8160	.7193	.6104	.4934	.3539	.2337	.1299
.9	.9811	.9431	.8764	.7741	.6639	.5496	.3999	.2688	.1499

Table 11.2.11. Power of T_4 -test with $\alpha = 0.05$.

$\rho \backslash p$	.1	.2	.3	.4	.5	.6	.7	.8	.9
.1	.1231	.1190	.1095	.0855	.0894	.0667	.0743	.0615	.0530
.2	.2290	.2102	.1842	.1625	.1285	.1097	.0924	.0866	.0637
.3	.3677	.3137	.2705	.2446	.2021	.1633	.1252	.0958	.0726
.4	.5005	.4370	.3793	.3178	.2651	.2146	.1727	.1207	.0769
.5	.6036	.5346	.4683	.3983	.3286	.2752	.2119	.1401	.0983
.6	.7140	.6426	.5533	.4865	.4064	.3178	.2431	.1648	.0986
.7	.7939	.7299	.6449	.5581	.4689	.3783	.2853	.1977	.1229
.8	.8552	.7836	.7191	.6277	.5304	.4290	.3302	.2257	.1288
.9	.9062	.8528	.7775	.6795	.5949	.4691	.3652	.2403	.1317

11.3. Diagrams

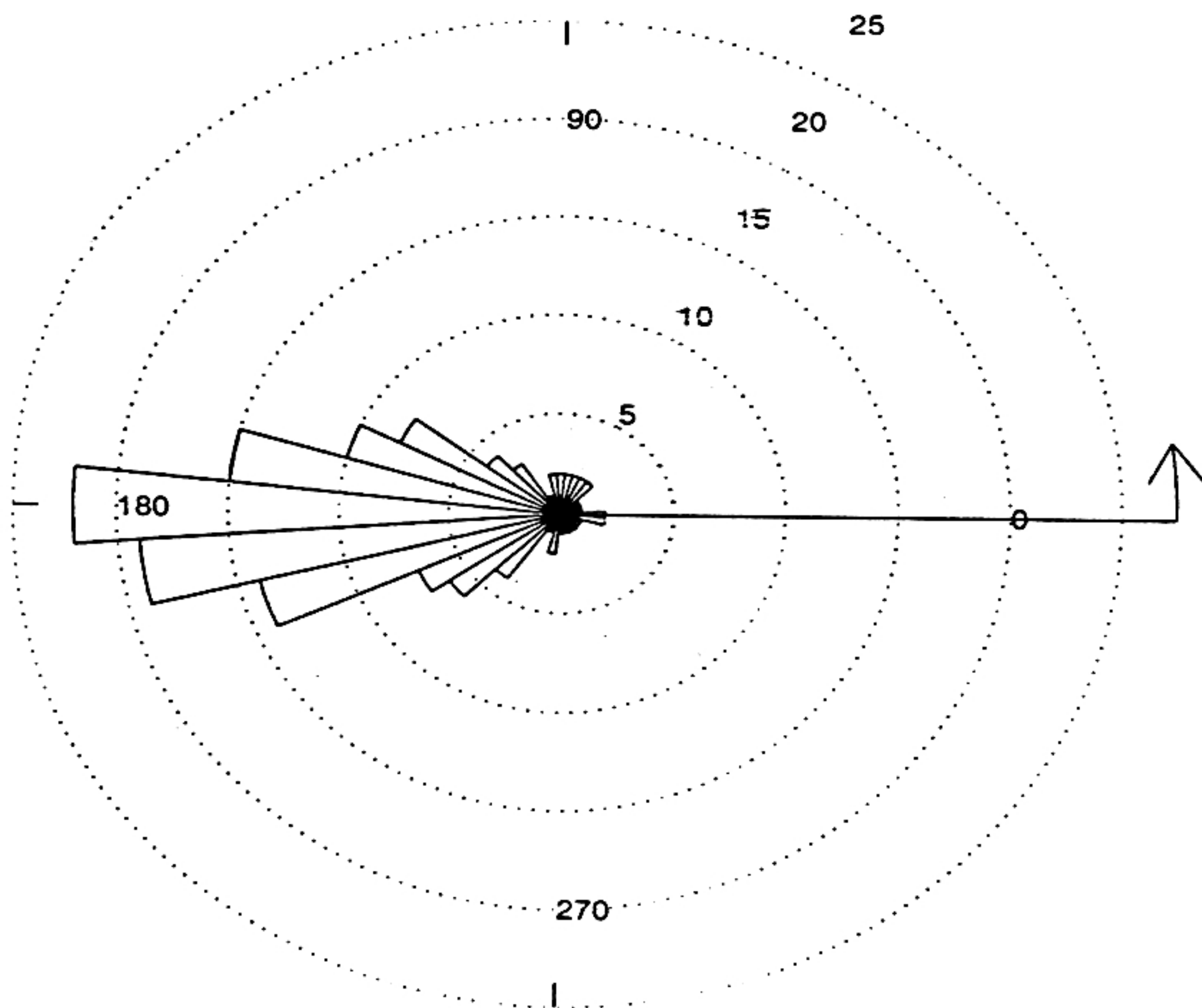


Figure 11.3.1. Rose diagram for Ant data

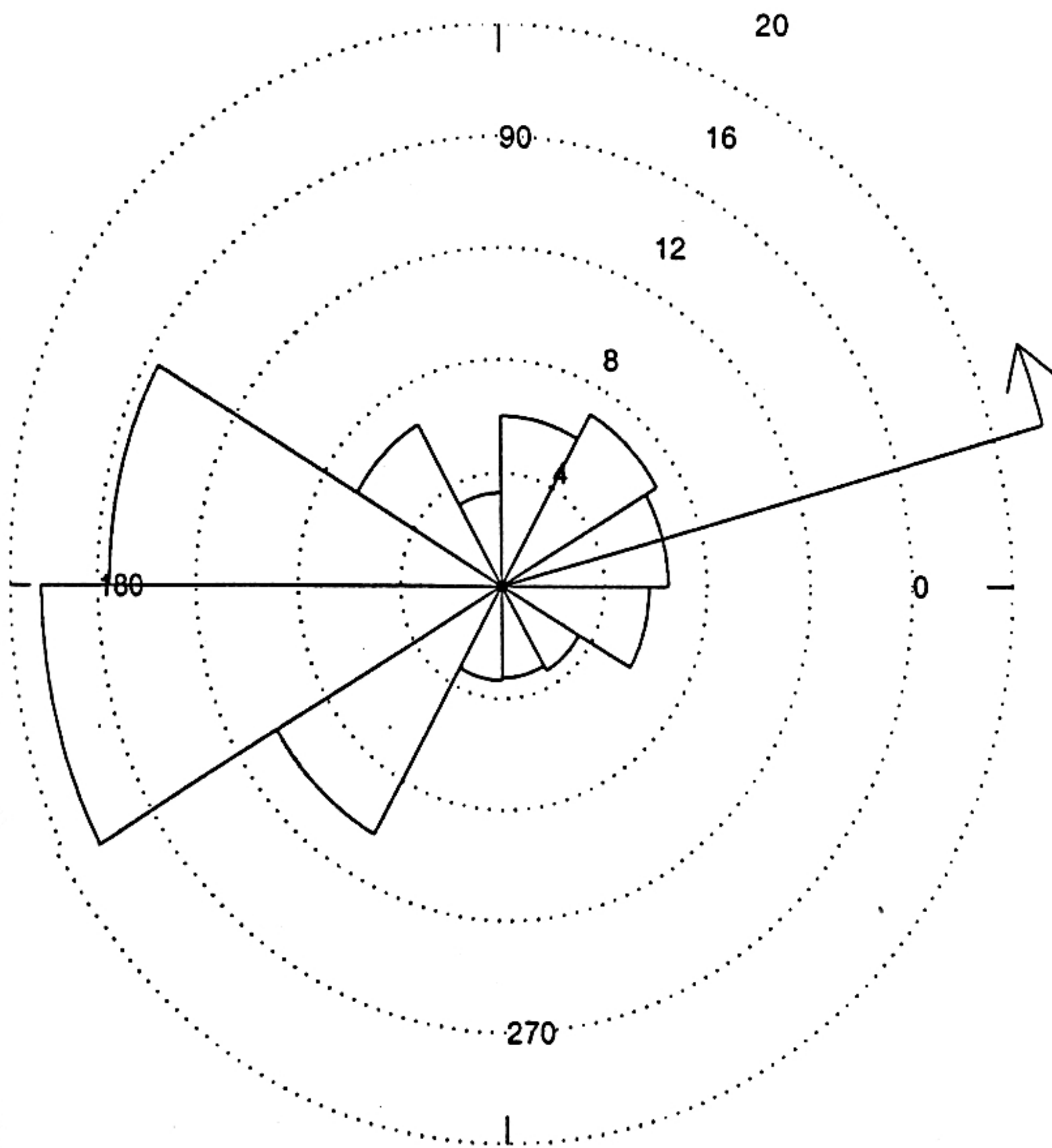


Figure 11.3.2. Rose diagram for Solar Wind data

Appendix A

Some Large Sample Results

Result A1. (Rao, 1973, pp. 387) Let T_n be a k -dimensional statistic $(T_{1n}, \dots, T_{kn})$ such that the a.d. of $(\sqrt{n}(T_{1n} - \theta_1), \dots, \sqrt{n}(T_{kn} - \theta_k))$ is k -variate normal with mean $\mathbf{0}$ and dispersion matrix $\Sigma = (\sigma_{ij})$. Further let g be a function of k variables which is totally differentiable. Then the a.d. of

$$\sqrt{n}u_n = \sqrt{n}[g(T_{1n}, \dots, T_{kn}) - g(\theta_1, \dots, \theta_k)]$$

is normal with mean zero and variance

$$v(\theta) = \sum \sum \sigma_{ij} \frac{\partial g}{\partial \theta_i} \frac{\partial g}{\partial \theta_j}.$$

Result A2. δ method (Rao, 1973, p. 388) : If g be a function as in Result A1 then the total differential of g is

$$dg = \frac{\partial g}{\partial \theta_1} dT_{1n} + \dots + \frac{\partial g}{\partial \theta_k} dT_{kn}$$

and assuming that the variances and covariances of $(T_{in} - \theta_i)$ exist,

$$V(dg) = \frac{1}{n} \sum \sum \frac{\partial g}{\partial \theta_r} \frac{\partial g}{\partial \theta_s} \sigma_{rs}.$$

Result A3. Slutsky's Theorem (Dudewicz and Mishra, 1988. p. 323) : Let $X_1, X_2, \dots$ be rv's and let $t_1 = t_1(X_1, \dots, X_n), \dots, t_r = t_r(X_1, \dots, X_n)$ be r functions of the rv's. Suppose that $t_1 \xrightarrow{P} a_1, \dots, t_r \xrightarrow{P} a_r$ (as $n \rightarrow \infty$) for some constants $a_1, \dots, a_r$. Suppose $g(\cdot, \dots, \cdot)$ is a rational function such that $|g(a_1, \dots, a_r)| < \infty$. Then

$$g(t_1, \dots, t_r) \xrightarrow{P} g(a_1, \dots, a_r).$$

Appendix B

Some Results on Fourier Analysis

Result B1. (Wintner, 1947, p. 591) The function $1 + 2 \sum_{n=1}^{\infty} \rho^n \cos n\theta$ is decreasing in $0 \leq \theta < \pi$ and by symmetry, increasing in $\pi \leq \theta < 2\pi$ irrespective of $0 < \rho < 1$ as long as $0 < a \leq 2$.

Result B2. (Wintner, 1947, p. 591) For $0 < \delta < \pi$, a set of sufficient conditions for the series $\sum_{n=1}^{\infty} b_n \sin n\delta$ to be positive is

- (i) $nb_n \rightarrow 0$ as $n \rightarrow \infty$ and
- (ii) $nb_n > (n+1)b_{n+1} \ \forall \ n = 1, 2, \dots$

Appendix C

Some Results on Stochastic Ordering

Result C1. (Lehmann, 1986, p. 434) Let X_λ denote a random variable distributed as noncentral χ^2 with f degrees of freedom and noncentrality parameter λ^2 . Then $X_{\lambda'}$ is stochastically larger than X_λ if $\lambda < \lambda'$.

The above Result is equivalent to the fact that the tail probability of a noncentral χ^2 variable is an increasing function of its noncentrality parameter.

Result C2. (SenGupta and Chang, 2000) Let (X, Y) have a joint absolutely continuous distribution depending on a parameter θ such that each of X and Y has the same stochastic ordering property (either increasing or decreasing) in θ . Then $X + Y$ has also the same stochastic ordering property.

This result can be easily generalized to any number of variables $X, Y, \dots$

Appendix D

Some Results on Riemann and Lebesgue Integrals

Result D1. (Apostol, 1974, Theorem 7.40, p. 167)

Let $Q = \{(x, y) : a \leq x \leq b, c \leq y \leq d\}$. Assume that α is of bounded variation on $[a, b]$ and, for each fixed y in $[c, d]$, assume that the integral

$$F(y) = \int_a^b f(x, y) d\alpha(x)$$

exists. If the partial derivative $D_2 f$ is continuous on Q , the derivative $F'(y)$ exists for each y in (c, d) and is given by

$$F'(y) = \int_a^b D_2 f(x, y) d\alpha(x).$$

Result D2. (Apostol, 1974, Theorem 10.26, p. 269) Let $\{g_n\}$ be a sequence of functions in $L(I)$ such that the series

$$\sum_{n=1}^{\infty} \int_I |g_n|$$

is convergent. Then the series $\sum_{n=1}^{\infty} g_n$ converges almost everywhere on I to a sum function g in $L(I)$ and we have

$$\int_I \sum_{n=1}^{\infty} g_n = \sum_{n=1}^{\infty} \int_I g_n.$$

Appendix E

Equicontinuity of a Family of Functions

Definition E1. (Dugundji, 1975, p. 266) Let (Z, d) be a metric space and Y be an arbitrary space. A subset $\mathcal{F} \subset Z^Y$ is equicontinuous at $y_0 \in Y$ if

$$\forall \epsilon > 0 \exists U(y_0) \forall f \in \mathcal{F} : f(U(y_0)) \subset B(f(y_0), \epsilon).$$

$U(y_0)$ is called a nbd of ϵ -equicontinuity for $\mathcal{F}$. We say that $\mathcal{F}$ is equicontinuous on Y whenever $\mathcal{F}$ is equicontinuous at each point of Y .

As an example, suppose $Y = [a, b]$, $Z = E^1$ and

$$\mathcal{F} = \{f \in Z^Y : |f'(y)| \leq M \text{ on } (a, b)\}.$$

Then $\mathcal{F}$ is equicontinuous on $[a, b]$.

Appendix F

Some Useful Results on Mixtures Based on Expansions

Result F1. (Durairajan, 1980, p. 148) Let $\alpha, \beta, x_1, \dots, x_n, y_1, y_2, \dots, y_n$ be real numbers. Then

$$\prod_{i=1}^n [\alpha x_i + \beta y_i] = \sum_{r=0}^n \sum_{C_r} \alpha^r \beta^{n-r} \prod_{i \in I_r} x_i \prod_{i \in I_r^c} y_i.$$

where for $r = 1, 2, \dots, n$, I_r denotes a particular combination of r integers $(i_1, i_2, \dots, i_r)$ out of $(1, 2, \dots, n)$, I_0 denotes a null set and C_r denotes the set of all such $\binom{n}{r}$ combinations.

Result F2. (Behboodian, 1972, p. 920) Let $X_1, X_2, \dots, X_n$ be a random sample from a mixture given by

$$f(x) = pf_1(x) + qf_2(x) \quad (0 < p < 1, \quad q = 1 - p)$$

and let $T = g(X_1, X_2, \dots, X_n)$ be a symmetric statistic. Let $T_k = g(X_{k1}, X_{k2}, \dots, X_{kn})$, $k = 0, 1, \dots, n$ be a statistic for which the X_{ki} 's are independent with density $f_1(x)$ if $i \leq k$ and density $f_2(x)$ if $i > k$. Then we have

$$f_T(t) = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} f_{T_k}(t),$$

that is the density of T is a binomial mixture of the densities of the T_k 's.

Appendix G

A Useful Result on the Exponential Family of Distributions

Result G1. (Lehmann, 1986, Theorem 9, p. 59) Let ϕ be any function on $(\mathcal{X}, \mathcal{A})$ for which the integral

$$\int \phi(x) \exp \left[\sum_{j=1}^s \theta_j T_j(x) \right] d\mu(x) \quad (1)$$

considered as a function of the complex variables $\theta_j = \xi_j + i\eta_j$ ($j = 1, \dots, k$) exists for all $(\xi_1, \dots, \xi_k) \in \Omega$ and is finite. Then

(i) the integral is an analytic function of each of the θ 's in the region R of parameter points for which $(\xi_1, \dots, \xi_k)$ is an interior point of the natural parameter space Ω ;

(ii) the derivatives of all orders with respect to the θ 's of the integral (1) can be computed under the integral sign.

BIBLIOGRAPHY

- Aitkin, M. and Rubin, D. B. (1985). Estimation and hypothesis testing in finite mixture models, *J. Roy. Statist. Soc., Ser. B*, **47**, 67-75.
- Apostol, T. (1974). *Mathematical Analysis*, 2nd ed., Addison Wesley.
- Arnold, B. C. and Strauss, D. (1988). Bivariate distributions with exponential conditionals, *J. Amer. Statist. Assoc.*, **83**, 522-527.
- Bacon-Shone, J. and Fung, W. K. (1987). A new graphical method for detecting single and multiple outliers in univariate and multivariate data, *App. Statist.*, **36**, 153-162.
- Bahadur, R. R. (1960). Stochastic comparison of tests, *Ann. Math. Statist.*, **31**, 276-295.
- Bahadur, R. R. (1967). Rates of convergence of estimates and test statistics, *Ann. Math. Statist.*, **38**, 303-324.
- Bahadur, R. R. and Raghavachari, M. (1972). Some asymptotic properties of likelihood ratio on general sample spaces, *Proc. Sixth Berkeley Symp. Math. Statist. Probability*, **1**, Univ. California Press, 129-152.
- Bai, Z. N., Chen, X. R., Miao, B. Q. and Rao, C. R. (1990). Asymptotic theory of least distances estimate in multivariate linear models, *Statistics*, **21**, 503-519.
- Barnett, V. and Lewis, T. (1994). *Outliers in Statistical Data*, 3rd ed., John Wiley & Sons.
- Bartels, R. (1984). Estimation in a bidirectional mixture of von Mises distributions, *Biometrics*, **40**, 777-784.
- Bartlett, M. S. and Please, N. W. (1963). Discrimination in the case of zero mean differences, *Biometrika*, **50**, 17-21.
- Batschelet, E. (1981). *Circular Statistics in Biology*, Academic Press.

- Behboodian, J. (1972). On the distribution of a symmetric statistic from a mixed population, *Technometrics*, **14**, 919-923.
- Beran, R. J. (1968). Testing for uniformity on a compact homogeneous space, *J. App. Probab.*, **5**, 177-195.
- Beran, R. J. (1969). Asymptotic theory of a class of tests for uniformity of a circular distribution, *Ann. Math. Statist.*, **40**, 1196-1206.
- Berdai, A. and Garel, B. (1996). Detecting a univariate normal mixture with two components, *Statist. Decisions*, **14**, 35-51.
- Bhattacharyya, G. K. and Johnson, R. A. (1969). On Hodge's bivariate sign test for uniformity of a circular distribution, *Biometrika*, **56**, 446-449.
- Bhatti, M. I. and King, M. L. (1990). A Beta-optimal test of the equicorrelation coefficient, *Austral. J. Statist.*, **32**, 87-97.
- Block, H. and Joe, H. (1997). Tail behavior of the failure rate functions of mixtures, *Lifetime Data Analysis*, **3**, 269-288.
- Box, G. E. P. and Tiao, G. C. (1968). A Bayesian approach to some outlier problems, *Biometrika*, **55**, 119-129.
- Brewer, K. R. W. and Tam, S. M. (1990). Is the assumption of uniform intra-class correlation ever needed? *Austral. J. Statist.*, **32**, 411-423.
- Brown, B. M. (1994). Grouping corrections for circular goodness-of-fit tests, *J. Roy. Statist. Soc., Ser. B*, **56**, 275-283.
- Buehler, W. J. and Puri, P. S. (1966). On optimal asymptotic tests of composite hypotheses with several constraints, *Z. Wahrsch. Verw. Gebiete*, **5**, 71-88.
- Butler, R. W. (1986). Predictive likelihood inference with applications (with discussion), *J. Roy. Statist. Soc. Ser. B*, **48**, 1-38.
- Cabral, M. S. E. (1991). A locally most powerful unbiased test for the proportion of mixture, *Statist. Prob. Letters*, **12**, 119-124.
- Chang, H. (1991). Some optimal tests in directional data, *Ph. D. dissertation, Dept. of Statistics and AP, UCSB, U.S.A.*
- Chen, J. (1994). Generalized likelihood-ratio test of the number of components in finite mixture models, *Canad. J. Statist.*, **22**, 387-399.

- Chernoff, H. and Lander, E. (1995). Asymptotic distribution of the likelihood ratio test that a mixture of two binomials is a single binomial, *J. Statist. Plann. Inference*, **43**, 19-40.
- Collett, D. (1980). Outliers in circular data, *App. Statist.*, **29**, 50-57.
- Cox, D. R. and Hinkley, D. V. (1974). *Theoretical Statistics*, Chapman and Hall, London.
- Cramer, H. (1946). *Mathematical Methods of Statistics*, Princeton University Press.
- D'Agostino, R. B. and Pearson, E. S. (1973). Tests for departure from normality. Empirical results for the distributions of b_2 and $\sqrt{b_1}$, *Biometrika*, **60**, 613-622.
- D'Agostino, R. B. and Pearson, E. S. (1974). Correction to "Tests for departure from normality. Empirical results for the distributions of b_2 and $\sqrt{b_1}$ ", *Biometrika*, **61**, 647.
- Davies, R. B. (1969). Beta-optimal tests and an application to the summary evaluation of experiments, *J. Roy. Statist. Soc. Ser. B*, **31**, 524-538.
- Davies, R. B. (1977). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, **64**, 247-254.
- Davies, R. B. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, **74**, 33-43.
- Dudewicz, E. Z. and Mishra, S. N. (1988). *Modern Mathematical Statistics*, John Wiley and Sons.
- Dugundji, J. (1975). *Topology*, Prentice Hall, New Delhi.
- Durairajan, T. M. (1980). Tests for the parameters of mixtures of distributions, *Ph. D. Thesis, University of Poona*.
- Durairajan, T. M. and Kale, B. K. (1979). Locally most powerful test for the mixing proportion, *Sankhya, Ser. B*, **41**, 91-100.
- Durairajan, T. M. and Kale, B. K. (1982). Locally most powerful similar test for mixing proportion, *Sankhya, Ser. A*, **44**, 153-161.

- Durairajan, T. M. and Raman, K. J. (1999). LBI tests for the detection of general compound distributions, *Cal. Statist. Assoc. Bull.*, **49**, 95-100.
- Eaton, M. L. and Perlman, M. D. (1973). The non-singularity of generalized sample covariance matrices, *Ann. Statist.*, **1**, 710-717.
- Efron, B. (1975). Defining the curvature of a statistical problem, *Ann. Statist.*, **3**, 1189-1242.
- Elston, R. C., Namboodiri, K. K., Nino, H. V. and Pollizer, W. S. (1974). Studies on blood and urine glucose in Seminole Indians: Indications for segregation of a major gene. *Amer. J. Human Genetics*, **26**, 13-34.
- Everitt, B. S. (1985). Mixture distributions, *Encyclopedia of Statistical Sciences* (eds. N. L. Johnson and S. Kotz), **5**, John Wiley, 559-569.
- Everitt, B. S. and Hand, D. J. (1981). *Finite Mixture Distributions*, Chapman and Hall.
- Feng, Z. and McCulloch, C. E. (1992). Statistical inference using maximum likelihood estimation and the generalized likelihood ratio when the true parameter is on the boundary of the parameter space, *Statist. Prob. Letters*, **13**, 325-332.
- Ferguson, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach*, Academic Press, New York.
- Furman, W. D. and Lindsay, B. G. (1994). Testing for the number of components in a mixture of normal distributions using moment estimators, *Comput. Statist. Data Anal.*, **17**, 473-492.
- Ganesalingam, S. and McLachlan, G. J. (1981). Some efficiency results for the estimation of the mixing proportion in a mixture of two normal distributions, *Biometrics*, **37**, 23-33.
- Ghosh, J. K. and Sen, P. K. (1985). On the asymptotic performance of the log likelihood ratio statistic for the mixture model and related results, In *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer, II* (eds. L. M. LeCam and R. A. Olshen), Wadsworth, Belmont, CA, 789-806.
- Gine, E. (1975). Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms, *Ann. Statist.*, **3**, 1243-1266.

- Goffinet, B., Loisel, P. and Laurent, B. (1992). Testing in normal mixture models when the proportions are known, *Biometrika*, **79**, 842-846.
- Gokhale, D. V. and SenGupta, A. (1986). Optimal tests for the correlation coefficient in a symmetric multivariate normal population, *J. Statist. Plann. Inference*, **14**, 263-268.
- Gunst, R. F. and Webster, J. T. (1973). Density functions of the bivariate chisquare distribution, *J. Statist. Comp. Simul.*, **2**, 275-288.
- Guttorp, P. and Lockhart, R. A. (1988). Finding the location of a signal: a Bayesian analysis, *J. Amer. Statist. Assoc.*, **83**, 322-330.
- Hajek, J. and Sidak, Z. (1967). *Theory of Rank Tests*, Academic Press.
- Hall, W. J. and Mathiason, D. J. (1990). On large-sample estimation and testing in parametric models, *Int. Statist. Rev.*, **58**, 77-97.
- Han, C. P. (1968). A note on discrimination in the case of unequal covariance matrices, *Biometrika*, **55**, 586-587.
- Hartigan, J. A. (1985). A failure of likelihood asymptotics for normal mixtures, In *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer, II* (eds. L. M. LeCam and R. A. Olshen), Wadsworth, Belmont, CA, 807-810.
- Henze, N. and Zirkler, B. (1990). A class of invariant consistent tests for multivariate normality, *Commun. Statist. – Theor. Meth.*, **19**, 3595-3617.
- Hsieh, H. K. (1979). On asymptotic optimality of likelihood ratio tests for multivariate normal distributions, *Ann. Statist.*, **7**, 592-598.
- Isaacson, S. L. (1951). On the theory of unbiased tests of simple statistical hypothesis specifying the values of two or more parameters, *Ann. Math. Statist.*, **22**, 217-234.
- Jander, C. (1957). Die optische Richtungsorientierung der roten Waldameise, (*Formica rufa* L.), *Z. vergl. Physiologie*, **40**, 162-238.
- Jiang, R. and Murthy, D. N. P. (1995). Modelling failure data by mixture of 2 Weibull distributions: A graphical approach, *IEEE Trans. Reliability*, **44**, 477-488.

- John, S. (1971). Some optimal multivariate tests, *Biometrika*, 58, 123-127.
- Johnson, N. L. and Kotz, S. (1970). *Continuous Univariate Distributions - 1*, John Wiley and Sons.
- Johnson, N. L. and Leone, F. C. (1977). *Statistics and Experimental Design in Engineering and the Physical Sciences*, vol 1, 2nd ed., Wiley, New York.
- Johnson, R. A. and Wichern, D. W. (1982). *Applied Multivariate Statistical Analysis*, Prentice Hall, New Jersey.
- Johnson, R. A. and Wichern, D. W. (1996). *Applied Multivariate Statistical Analysis*, 3rd ed., Prentice Hall of India.
- Jupp, P. E. and Mardia, K. V. (1989). A unified view of the theory of directional statistics, *Int. Statist. Rev.*, 57, 261-294.
- King, M. L. (1987). Towards a theory of point optimal testing, *Econometric Rev.*, 6, 169-218.
- Koch, G. C. (1983). Intraclass correlation coefficient, *Encyclopedia of Statistical Sciences* (eds. N. L. Johnson and S. Kotz), 4, 212-217.
- Koch, G. C., Elashoff, J. D. and Amar, I. A. (1988). Repeated measurements – design and analysis, *Encyclopedia of Statistical Sciences*, eds. N. L. Johnson and S. Kotz, 8, 46-73.
- Kocherlakota, K. and Kocherlakota, S. (1981). On the distribution of r in samples from the mixtures of bivariate normal populations, *Commun. Statist. – Theor. Meth.*, 10, 1943-1966.
- Kocherlakota, K. and Kocherlakota, S. (1985). On some tests for independence in nonnormal situations : Neyman's $C(\alpha)$ test, *Commun. Statist. – Theor. Meth.*, 14, 1453-1470.
- Kocherlakota, S. (1986). The bivariate inverse Gaussian distribution: An introduction, *Commun. Statist. – Theor. Meth.*, 15, 1081-1112.
- Lehmann, E. L. (1983). *Theory of Point Estimation*, John Wiley and Sons.
- Lehmann, E. L. (1986). *Testing Statistical Hypotheses*, 2nd ed., John Wiley and Sons.

- Lehmann, E. L. and Casella, G. (1998). *Theory of Point Estimation*, 2nd ed., Springer.
- Lockhart, R. A. and Stephens, M. A. (1985). Tests of fit for the von Mises distribution, *Biometrika*, **72**, 647-652.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications, *Biometrika*, **57**, 519-530.
- Mardia, K. V. (1972). *Statistics of Directional Data*, Academic Press.
- Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies, *Sankhya, Ser. B*, **36**, 115-128.
- Mardia, K. V. and Zemroch, P. J. (1975). Measures of multivariate skewness and kurtosis, *App. Statist.*, **24**, 262-265.
- Marks, R. G. and Rao, P. V. (1979). An estimation procedure for data containing outliers with a one-directional shift in the mean, *J. Amer. Statist. Assoc.*, **74**, 614-620.
- McLachlan, G. J. (1992). *Discriminant Analysis and Statistical Pattern Recognition*, Wiley, New York.
- McLachlan, G. J. (1997). *Recent Advances in Finite Mixture Models*, John Wiley, New York.
- McLachlan, G. J. and Basford, K. E. (1988). *Mixture Models : Inference and Applications to Clustering*, Marcel Dekker, New York.
- Mendell, N. R., Finch, S. J. and Thode, H. C. (Jr.) (1993). Where is the likelihood ratio test powerful for detecting two-component normal mixtures? *Biometrics*, **49**, 907-915.
- Moran, P. A. P. (1970). On asymptotically optimal tests of composite hypotheses, *Biometrika*, **57**, 47-55.
- Mukherjee, S. P. and Chatterjee, A. (1987). Estimation of parameters in finite mixtures of exponential families from censored data, *Commun. Statist. - Theor. Meth.*, **16**, 3133-3147.

- Newcomb, S. (1886). A generalized theory of the combinations of observations so as to obtain the best result, *Amer. J. Math.*, **8**, 343-366.
- Neyman, J. (1959). Optimal asymptotic tests of composite statistical hypotheses, In *Probability and Statistics, the Harald Cramer Volume* (ed. U. Grenander), Almqvist and Wiksell, Uppsala, Sweden, 213-234.
- Neyman, J. and Pearson, E. S. (1938). Contributions to the theory of testing statistical hypothesis, *Statist. Res. Memoirs*, **2**, 25-57.
- Odell, P. L. (1976). Current and past roles of the statistician in space applications, *Commun. Statist.- Theor. Meth.*, **5**, 1091-1114.
- Pal, C. (2000a). On a locally Bahadur optimal test for no contamination in mixtures from the one-parameter exponential family, *Submitted for publication*.
- Pal, C. (2000b). Higher order $C(\alpha)$ tests with applications to mixture models, *Submitted for publication*.
- Pal, C. and SenGupta, A. (2000a). Optimal tests for no contamination in reliability models, *Lifetime Data Analysis*, **6**, 281-290.
- Pal, C. and SenGupta, A. (2000b). On P^3 -tests in multivariate mixtures with applications to normal distributions, *Submitted for publication*.
- Pearson, E. S. and Hartley, H. O. (Eds.) (1972). *Biometrika Tables for Statisticians, vol. II*, Cambridge University Press.
- Pearson, K. (1894). Contribution to the mathematical theory of evolution, *Phil. Trans. Roy. Soc., Ser A*, **185**, 71-110.
- Philips, J. L. et al. (1995). Ulysses solar wind plasma observations at high southerly latitudes, *Science*, **268**, 1030-1033.
- Quinn, B. G., McLachlan, G. J. and Hjort, N. L. (1987). A note on the Aitkin-Rubin approach to hypothesis testing in mixture models, *J. Roy. Statist. Soc., Ser. B*, **49**, 311-314.
- Rachev, S. T. and SenGupta, A. (1992). Geometric stable distributions and Laplace-Weibull mixtures, *Statist. Decisions*, **10**, 251-271.

- Rao, C. R. (1973). *Linear Statistical Inference and its Applications*, 2nd ed., Wiley, New York.
- Roy, S. N. (1953). On a heuristic method of test construction and its use in multivariate analysis, *Ann. Math. Statist.*, **24**, 220-238.
- Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components, *Biometrics Bull.*, **2**, 110-114.
- Self, S. G. and Liang, K. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions, *J. Amer. Statist. Assoc.*, **82**, 605-610.
- SenGupta, A. (1982). On tests for equicorrelation coefficient and the generalized variance of a standard symmetric multivariate normal distribution, *Tech. Report #54, Department of Statistics, Stanford University, California*.
- SenGupta, A. (1987). On tests for equicorrelation coefficient of a standard symmetric multivariate normal distribution, *Austral. J. Statist.*, **29**, 49-59.
- SenGupta, A. (1991). A review of optimality of multivariate tests, In *Special issue on multivariate optimality and related topics* (guest eds. S. R. Jammalamadaka and A. SenGupta), *Statist. Prob. Letters*, **12**, 527-535.
- SenGupta, A. (1998a). Optimal parametric statistical inference, In *Analysis of Directional Data* (ed. A. SenGupta), Indian Statistical Institute, Calcutta, 23-52.
- SenGupta, A. (1998b). DDSTAP – Statistical package for the analysis of directional data, *Applied Statistics Unit, Indian Statistical Institute, Calcutta*.
- SenGupta, A. and Chang, H. (2000). A robust optimal test for uniformity for directional data against wrapped stable mixture family, *Submitted for publication*.
- SenGupta, A. and Pal, C. (1991). Locally optimal tests for no contamination in standard symmetric multivariate normal mixtures, *J. Statist. Plann. Inference*, **29**, 145-155.
- SenGupta, A. and Pal, C. (1993a). Optimal tests in some applied mixture models with non-regularity problems, In *Probability and Statistics* (eds. S. K. Basu and B. K. Sinha), Narosa Publishing House, New Delhi, 151-164.

- SenGupta, A. and Pal, C. (1993b). Optimal tests for no contamination in symmetric multivariate normal mixtures, *Ann. Inst. Statist. Math.*, **45**, 137-146.
- SenGupta, A. and Pal, C. (1995). On optimal tests for no contamination in reliability models, *Paper presented at National Seminar on Survival Analysis and Reliability*, University of Calcutta.
- SenGupta, A. and Pal, C. (2001). On optimal tests for isotropy against the symmetric wrapped stable - circular uniform mixture family, *J. App. Statist.*, **28**, 129-143 (To appear).
- SenGupta, A. and Vermeire, L. (1986). Locally optimal tests for multiparameter hypotheses, *J. Amer. Statist. Assoc.*, **81**, 819-825.
- Spjøtvoll, E. (1968). Most powerful tests for some non-exponential families, *Ann. Math. Statist.*, **39**, 772-784.
- Srivastava, M. S. and Lee, G. C. (1984). On the distribution of the correlation coefficient when sampling from a mixture of two bivariate normal densities: robustness and the effect of outliers, *Canad. J. Statist.*, **12**, 119-133.
- Sugiyama, T. (1972). Percentile points of the largest latent root of a matrix and power calculations for testing $\Sigma = I$, *J. Jap. Statist. Soc.*, **3**, 1-8.
- Titterington, D. M. (1997). Mixture distributions (update), *Encyclopedia of Statistical Sciences*, Update Vol 1 (eds. S. Kotz, C. B. Read and D. L. Banks), John Wiley, 399-407.
- Titterington, D. M., Smith, A. F. M. and Makov, U. E. (1985). *Statistical Analysis of Finite Mixture Distributions*, Wiley, New York.
- Verrill, S., Axelford, M. and Durst, M. (1990). We've got the positive correlation BLUEs, *The American Statistician*, **44**, 171-173.
- Wintner, A. (1947). On the shape of the angular case of Cauchy's distribution curves, *Ann. Math. Statist.*, **18**, 589-593.
- Yamamoto, W. and Shinozaki, N. (2000). On uniqueness of two principal points for univariate location mixtures, *Statist. Prob. Letters*, **46**, 33-42.

INDIAN STATISTICAL INSTITUTE LIBRARY

Processing slip

Source.....P/E/G date. 23.12.65 Acc. no. T152

Author... Pal Chandanath

Title... On some optimal tests in finite mixture

CHECKING :

Call no.

1 In library/earlier ed./later ed.

2 Earlier volumes/nos. of
the series in lib.

3 New title ☒

Suggested/Approved by

Note :

23.12.65
Signature & date

CLASSIFICATION :

Class no.....x-ref.....

Scrutiny report :

Signature & date

Class no.....x-ref.....

Note :

Signature & date

CATALOGUING :

1 Main Card

2 Supp. Cards (Subject/Title/Series/Added entries/Shelf List) [Total no. of
cards].....

3 Mechanical processing completed on.....

4 Cards filed

Note :

Signature & date