Journal of Statistical Planning and Inference 29 (1991) 145-155
North-Holland

Locally optimal tests for no contamination in standard symmetric multivariate normal mixtures

Ashis SenGupta

Indian Statistical Institute, Calcutta, W.B. 700035, India, and University of California, Santa Barbara, CA 93106, USA

ChandraNath Pal*

Indian Statistical Institute, Calcutta, W.B. 700035, India

Received 22 March 1989; revised manuscript received 30 October 1989

Abstract: Mixture models are becoming increasingly popular in reliability studies and in survival analysis. Consider a mixture of two standard symmetric multivariate normal distributions with one of them having intraclass correlation coefficient, ϱ as 0. We are interested in testing for no contamination. To avoid the problem of identifiability, we assume that either $\varrho \neq 0$ or that the mixing proportion $p<1$ and then test for the other parameter. From practical and theoretical considerations based on statistical curvature, we advocate the test for H_0: $\varrho=0$ in preference to that for H_0: $p=1$. Serious complications still exist, so we concentrate on the problem when p is known. Assuming p known, the locally most powerful test is shown to be extremely simple, and the exact cut-off point is easily computed. The test is admissible, unbiased and possesses a globally monotone power function. Computation of the exact power values exhibit encouraging results. The test statistic is asymptotically normally distributed under both the null and alternative hypotheses. Finally, the test is shown to be also consistent.

AMS Subject Classification: Primary 62H10.

Key words and phrases: Intraclass correlation coefficient; locally most powerful test; mixture distribution; standard symmetric multivariate normal distribution; statistical curvature.

1. Introduction and summary

Mixture models are now increasingly playing important and popular roles in reliability, see e.g. Titterington, Smith and Makov (1985, p. 17 and p. 20). Symmetric multivariate normal (SMN) distribution with equicorrelation pattern (Rao, 1973) for the correlation matrix is also a very popular model in reliability studies.

* Currently at University of Kalyani, India.

However, little seems to be known about any method of construction of optimal tests in mixture models. Here we consider the test that the underlying population is standard SMN (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: $\varrho=0$, H_0: $p=1$ and H_0: $\varrho=0$, $p=1$ simultaneously, leads to the hypothesis of no mixture. So, to avoid the problem of identifiability, we might assume either $p\neq 1$ or $\varrho\neq 0$. 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 likelihood ratio test (LRT) statistic is not even available in a closed form. A locally most powerful (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 $\varrho=0$ with variations of p in $(0,1)$. Also, the test statistic simply coincides with the one for H_0: $\varrho=0$ 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: $\varrho=1$ 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: $\varrho=0$ is justified in favor of that for H_0: $p=1$.

The proposed LMP test is admissible. 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 $p=0.9$. The test has a monotone power function for alternatives, $\varrho>0$, is locally consistent and the test statistic is asymptotically normally distributed under both the null and the alternative hypotheses. This is 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 likelihood ratio test statistic, λ, is quite complicated.

2. Locally most powerful test

Let $g(\boldsymbol{x} \mid m, \sigma^2, \varrho)$ denote the k-variate SMN density (Rao, 1973, p. 196), $\mathrm{N}_k(\boldsymbol{M}, \sigma^2 \Sigma_\varrho)$, with mean vector $\boldsymbol{M}=(m, \ldots, m)'$, $\sigma^2>0$, $\Sigma_\varrho=((\varrho+(1-\varrho)\delta_{ij}))$, δ_{ij} being the Kronecker delta and $-(k-1)^{-1}<\varrho<1$. This density can also be regarded as a density of k exchangeable normal random variables with the same marginal parameters m and σ^2. Let $g_{(p)}(\boldsymbol{x} \mid m, \sigma^2, \varrho) \equiv g_{(p)}$ be the density obtained as a 'p-mixture' of $g(\boldsymbol{x} \mid m, \sigma^2, 0) \equiv g_0$ and $g(\boldsymbol{x} \mid m, \sigma^2, \varrho) \equiv g_\varrho$, that is

$$
\begin{aligned}
g_{(p)}(\boldsymbol{x} \mid m, \sigma^2, \varrho) = {} & p(2\pi\sigma^2)^{-k/2} \exp\{-(\boldsymbol{x}-\boldsymbol{M})'(\boldsymbol{x}-\boldsymbol{M})/2\sigma^2\} \\
& + q(2\pi\sigma^2)^{-k/2}(\det \Sigma_\varrho)^{-1/2} \exp\{-(\boldsymbol{x}-\boldsymbol{M})'\Sigma_\varrho^{-1}(\boldsymbol{x}-\boldsymbol{M})/2\sigma^2\} \quad (1.1)
\end{aligned}
$$

where $0<p<1$ and $q=1-p$. Let $X_1, \ldots, X_n$ be a random sample from (1.1). Assume $m=0$, $\sigma^2=1$. Here we derive LMP test of H_0: $\varrho=0$ against the one sided alternatives H_1: $\varrho>0$. Note that for large k, ϱ should be non-negative. Usual reversal of inequality in the definition of the critical region yields an analogous test for the alternatives $\varrho<0$. 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 LMP test.

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

$$
\sum_{s=1}^{n} \left[\frac{\partial}{\partial \varrho} \ln g_{(p)}(\boldsymbol{x}_s \mid m, \sigma^2, \varrho) \right] \Bigg|_{\varrho=0} > c \tag{2.1}
$$

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

Since p is known, (2.1) reduces to a similar expression with $g_{(p)}$ being replaced by g_ϱ, 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 $\boldsymbol{X}_s=(X_{s1}, X_{s2}, \ldots, X_{sk})'$, X_{si} denoting the i-th component of the vector $\boldsymbol{X}_s$, $\bar{\boldsymbol{X}}$ denote the sample mean vector, $\bar{X}=(\sum_{s=1}^n \sum_{i=1}^k X_{si})/nk$, and $\bar{X}_s$ be defined by $\bar{X}_s=(\sum_{i=1}^k X_{si})/k$, $s=1,2,\ldots,n$. Let

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

and

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

Under g_ϱ, $W/(1-\varrho)$ is distributed as $\chi^2_{n(k-1)}$, $B/\{1+(k-1)\varrho\}$ is distributed as χ^2_{n-1} and W and B are independent (Rao, 1973). 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.

3. Statistical curvature

3.1. Definition

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 $\gamma_{\theta_0}^2 < \frac{1}{8}$ 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 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 $\frac{1}{8}$. Let $\mathscr{F} = \{f_\theta(x), \theta \in \Theta\}$ be an arbitrary family of one-parameter density functions. Let, $l_\theta(x) = \ln f_\theta(x)$. Then, under the usual assumptions of existence of derivatives etc., the statistical curvature of f at θ is defined as γ_θ,

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

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: $\varrho = 0$, 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.

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

Let $l_q \equiv \ln g_{(p)}$. Then $\dot{l}_0 = r - 1$ and $\ddot{l}_0 = -(r-1)^2$ where $r \equiv g_\varrho / g_0$. Now,

$$r = |\Sigma_\varrho|^{-1/2} \exp[\{-\varrho/2k(1-\varrho)(1+(k-1)\varrho)\} \times \{\textstyle\sum (x_i - \bar{x})^2 - (k-1)(1-\varrho)(\sum x_i)^2\}] \tag{3.1}$$

$$\beta_j \equiv E_{q=0}(r^j) = (2\pi)^{-k/2} |\Sigma_\varrho|^{-j/2} \int \exp\left[-\frac{1}{2}\left\{\frac{(j-1)\varrho + 1}{1-\varrho} \sum x_i^2 - \frac{j\varrho}{(1-\varrho)(1+(k-1)\varrho)} (\textstyle\sum x_i)^2\right\}\right] \mathrm{d}x \tag{3.2}$$

$$= (2\pi)^{-k/2} |\Sigma_\varrho|^{-j/2} \int \exp[-\tfrac{1}{2} x' B_j x] \, \mathrm{d}x, \quad \text{say,} \tag{3.3}$$

where,

$$x'B_jx = x'\left[\frac{(j-1)\varrho+1}{1-\varrho} I_k - \frac{j\varrho}{(1-\varrho)(1+(k-1)\varrho)} 11'\right]x.$$

Hence from (3.3), when defined and when it exists, β_j is given by

$$\beta_j = (1-\varrho)^{\{k-j(k-1)\}/2}\{1+(k-1)\varrho\}^{-j/2}$$
$$\times[\{(j-1)\varrho+1\}^k - \{(k\varrho j)((j-1)\varrho+1)^{k-1}/(1+(k-1)\varrho)\}]^{-1/2}. \quad (3.4)$$

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, ${}_1\gamma^2_{q=0} = (\alpha_4-\alpha_2^2)/\alpha_2^2 - \alpha_3^2/\alpha_2^3$.

3.3. ${}_1\gamma^2_{\varrho=0}$ as a function of q

Let $l_\varrho \equiv \ln g_{(p)}$. Then $\dot{l}_0 = q\dot{g}_0/g_0$ and $\ddot{l}_0 = (qg_0\ddot{g}_0 - q^2\dot{g}_0^2)/g_0^2$. Let

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

and

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

Under g_0, T_1 and T_2 are independently distributed with $T_1 \sim \chi^2_1$, $T_2 \sim \chi^2_{k-1}$. After some tedious calculations, and using the fact that if $Y \sim \chi^2_n$, then

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

we have

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

$$E(\dot{l}_0\ddot{l}_0) = \tfrac{1}{2}q^2 E[\{(k-1)T_1 - T_2\}\{\tfrac{1}{2}k(k-1) + \tfrac{1}{4}((k-1)T_1 - T_2)^2$$
$$- (T_2 + (k-1)^2T_1)\}] - \tfrac{1}{8}q^3 E\{(k-1)T_1 - T_2\}^3$$
$$= -q^3k(k-1)(k-2),$$

and

$$E(\ddot{l}_0^2) = q^2E(T_3^2) + \tfrac{1}{16}q^4 E\{(k-1)T_1 - T_2\}^4 - (q^{3/2})E[T_3\{(k-1)T_1 - T_2\}^2]$$
$$= q^2\{\tfrac{3}{4}k^4 - 5k^3 + \tfrac{13}{2}k^2 - 3k\} + \tfrac{1}{16}q^4(60k^4 - 216k^3 + 300k^2 - 144k)$$
$$- \tfrac{1}{2}q^3(6k^4 - 20k^3 + 26k^2 - 12k).$$

Thus,

$$|M_0| = \{\tfrac{1}{4}k^2(k-1)^2\}\{q^6(3k^2-3k+2) - 2q^5(3k^2-7k+6) + q^4(3k^2-7k+6)\}.$$

Finally

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

3.4. Comparison and comments

Let $\gamma^2(1) \equiv {}_1\gamma^2_{\varrho=0}$ and $\gamma^2(2) \equiv {}_1\gamma^2_{q=0}$. Consider $\gamma^2(2)$. Rewrite (3.2) in terms of

Table 3.1
Values of ${}_1\gamma^2{}_{\varrho=0}$ with $q=0.05(0.10)1.00$ and $k=2(2)10$

q	k: 2	4	6	8	10
0.05	1448.00	1566.33	1734.13	1831.79	1894.04
0.15	132.44	141.14	155.47	163.85	169.20
0.25	40.00	41.00	44.53	46.54	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

$\sum (x_i-\bar{x})^2$ and $(\sum x_i)^2$ as in (3.1). Then, the coefficient of $(\sum x_i)^2$ becomes negative for values of ϱ, k and j such that

$$(k-1)(j-1)\varrho^2+(k+j-kj-2)\varrho+1<0 \tag{3.5}$$

for example, with $\varrho=0.05$, $k=8$ and $j=4$. So, β_j may not even be defined for such values of ϱ, k and j which is seen from (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 $\gamma^2(1)\to 8/k=\gamma^2(1)\,|_{q=1}$ as $q\to 1$. We recall (SenGupta, 1987) that $8/k$ is the statistical curvature of the CEF g_ϱ at $\varrho=0$. One can show that, for all k, $\gamma^2(1)\downarrow k$ also, but with $q\geq 0.5$. $\gamma^2(1)\to\infty$ as $q\to 0$. Thus the LMP test for $\varrho=0$ can be expected to perform well for all k and for all q not too small. The conservative behavior for q small is to be naturally expected. The values of $\gamma^2(1)$ exhibited in Table 3.1 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 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: $\varrho=0$.

4. Exact power computations

We have

$$T_1=\sum_{s=1}^{n}\sum_{i\neq j=1}^{k}X_{si}X_{sj}=\sum_{s=1}^{n}\left[\left(\sum_{i=1}^{k}X_{si}\right)^2-\sum_{i=1}^{k}X_{si}^2\right]$$
$$=\sum_{s=1}^{n}\left[k(k-1)\bar{X}_s^2-\sum_{i=1}^{k}(X_{si}-\bar{X}_s)^2\right]. \tag{4.1}$$

Under g_0, $\bar{X}_s \sim \mathrm{N}(0, 1/k)$ and hence $k\bar{X}_s^2 \sim \chi_1^2$. Again, $\sum_{i=1}^{k} (X_{si} - \bar{X}_s)^2 \sim \chi_{k-1}^2$ under g_0 and $\sim (1-\varrho)\chi_{k-1}^2$ under g_ϱ. So (4.1) becomes $T_1 = \sum_{s=1}^{n} Y_s$, Y_s's are i.i.d.,

$$Y_s \sim \begin{cases} (k-1)\chi_1^2 - \chi_{k-1}^2 & \text{under } g_0, \\ (k-1)\{1+(k-1)\varrho\}\chi_1^2 - (1-\varrho)\chi_{k-1}^2 & \text{under } g_\varrho. \end{cases} \tag{4.2}$$

Hence, since T_1 is a symmetric statistic of Y_i's, we have

$$P(T_1 > c \mid \varrho) = 1 - \sum_{s=0}^{n} \binom{n}{s} p^s (1-p)^{n-s} P\left[\sum_{i=0}^{s} Y_i + \sum_{i=s+1}^{n} Y_i \leq c \right]$$

where $Y_0 \equiv 0$, Y_i, $i = 1, \ldots, s$, follows g_0 and Y_i, $i = s+1, \ldots, n$, follows g_ϱ. Thus,

$$\begin{aligned} P(T_1 > c \mid \varrho) = 1 - \sum_{s=0}^{n} \binom{n}{s} p^s (1-p)^{n-s} P[&(k-1)\chi_s^2 - \chi_{s(k-1)}^2 \\ &+ (k-1)\{1+(k-1)\varrho\}\chi_{(n-s)}^2 \\ &- (1-\varrho)\chi_{(n-s)(k-1)}^2 \leq c]. \end{aligned} \tag{4.3}$$

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 Table 2 of 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 χ_s^2, $\chi_{s(k-1)}^2$, $\chi_{(n-s)}^2$ and $\chi_{(n-s)(k-1)}^2$, $s = 0, 1, \ldots, n$, were generated with obvious modifications for $s=0$ and $s=n$. Then, the probability, inside the summation sign on the right hand side of (4.3), for each s, is obtained as the corresponding empirical c.d.f. So, for example, to obtain Table 4.1, the simulation loop was executed

Table 4.1
Exact power values for LMP test ($\alpha = 0.05$, $k=3$; $n=10$ and $n=25$) [a]

ϱ	p: 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

[a] First and second sets of entries correspond to $n=10$ and $n=25$ respectively.

$p\cdot\varrho\cdot(n+1)\cdot 1000\cdot 4=9\cdot 10\cdot 26\cdot 1000\cdot 4=9\,360\,000$ times. From the first set of entries of Table 4.1 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 which is also established analytically in Section 5 below. The rapid increase in power, for small p, with increase in n is noted from the second set of entries. For large p, the power is not high as is to be expected.

5. Unbiasedness and monotonicity

Note that the last expression in the right of (4.2) is an increasing function of ϱ for $\varrho\geq 0$. This implies that for $0\leq\varrho_1<\varrho_2\leq 1$, Y_s is (stochastically) larger under g_{ϱ_2} than under g_{ϱ_1}. From (4.3) it follows that, the power function is monotonically increasing and hence the test is globally unbiased against one-sided alternatives, $\varrho>0$.

6. Asymptotic distribution of the test statistic

Let $\tilde{\varrho}=T_1/nk(k-1)$. We show below that $\tilde{\varrho}$ is asymptotically normally distributed under both the null and the alternative hypotheses. This result is quite reassuring and important. Recall (Ghosh and Sen, 1985) that $-2\ln\lambda$ 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 $p\mathrm{N}(0,1)+q\mathrm{N}(\theta,1)$ where $-2\ln\lambda\xrightarrow{\mathrm{P}}\infty$. 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,

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

where

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

and $\chi^2_{\nu_1}$ and $\chi^2_{\nu_2}$ are independently distributed under g_ϱ. Then,

$$E\tilde{\varrho}=p0+q\varrho=q\varrho\quad\text{and}\quad \mathrm{Var}(\tilde{\varrho})=[EU^2-(q\varrho)^2]/n.$$

Under g_ϱ, we have

$$\begin{aligned}E_\varrho U^2&=3\alpha_1^2+(k^2-1)\alpha_2^2-2(k-1)\alpha_1\alpha_2\\&=[(3k^2-7k+6)k\varrho^2+4k(k-2)\varrho+2k]/k^2(k-1)\equiv B(\varrho),\quad\text{say.}\end{aligned}$$

Then, under $g_{(p)}$, we have

$$\begin{aligned}\mathrm{Var}(U) &= [pB(0)+qB(\varrho)]/k^2(k-1)-(q\varrho)^2 \\ &= [2kp+B(\varrho)q]/k^2(k-1)-(q\varrho)^2 \equiv \sigma^2(\varrho), \quad \text{say.} \end{aligned} \tag{6.1}$$

Then, by the central limit theorem, we have:

Theorem 1. *Under $g_{(p)}$, $(\tilde{\varrho}-q\varrho)$ is asymptotically distributed as a normal variable with mean* 0 *and variance $\sigma^2(\varrho)$ given in* (6.1).

Letting $q=1$ in the above theorem, we have, as a special case, Theorem 2 of SenGupta (1987), which gave the asymptotic distribution of $\tilde{\varrho}$ for the SSMN distribution.

7. Consistency of the LMPU test

Let α be the level of the test and c the cut-off point so that

$$\alpha = P[\tilde{\varrho} > c/n \mid \mathrm{H}_0].$$

By Theorem 1, then, $c \simeq \sqrt{n}\sigma(0)\tau_\alpha$. So,

$$\begin{aligned}\mathrm{Power}(\varrho) &= P[\tilde{\varrho} > c/n \mid \mathrm{H}_1] \\ &\simeq 1-\Phi[\{\sigma(0)\tau_\alpha - n^{1/2}q\varrho\}/\sigma(\varrho)] \quad \text{(by Theorem 1)} \\ &\to 1 \quad \text{as } n\to\infty, \quad \text{since } \varrho>0.\end{aligned}$$

Thus the LMPU test is globally ($\varrho>0$) consistent. We have already seen in Section 5 that this test is also globally unbiased and possesses a globally monotone power function.

8. Remarks

For the general problem of testing 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 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, $g_{(p)}(x)=pg_1(x,\theta)+qg_2(x,\eta)$. 'In most applications', however, as Titterington 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, $g_{(p)}(x)=pg(x,\theta_1)+qg(x,\theta_2)$, 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 $\theta_1\neq\theta_2$. 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, $\theta_1\equiv\varrho$, $\theta_{10}=0$; etc. Then the assumption of p known leads to the LMP test for H_0: no mixture in $g_{(p)}(x)$, to coincide with that for H_0: $\theta_1=\theta_{10}$ in $g(x,\theta)$. Thus, often the test statistic is quite simple and exact and elegant results regarding its properties may be obtained as in this paper. Also for this case of p known, even in the presence of nuisance parameters (e.g. m and σ^2 in our problem) one can proceed to construct LMP similar and/or invariant tests. Further, if p is also a nuisance parameter, one can attempt to construct asymptotically LMP tests, e.g. Neyman's c_α-test, likelihood ratio derivative test, etc. Finally, even if p is to be tested also, one can explore the multiparameter LMMPU tests of SenGupta and Vermeire (1986), e.g. to test, for our problem, H_0: $p=p_0$ and $\varrho=0$, $p\in(0,1)$, $\varrho\in(-1/(k-1),1)$, simultaneously. Our approach, in the given framework, is thus applicable quite generally and the above observations are interesting topics for further research.

Acknowledgements

The authors would like to thank the referee for helpful comments to improve the presentation of the paper. The first author thankfully acknowledges the financial grants from the conference committee and the Indian Statistical Institute which made possible his presentation of the paper at the conference. The authors are grateful to Professors A.P. Basu and J.K. Ghosh for their encouragements and helpful discussions. They would also like to thank Mr. S. Nandi, I.S.I. and Mr. C. Dawson, University of California, Santa Barbara, for the computations of the tables.

References

Butler, R.W. (1986). Predictive likelihood inference with applications (with discussion). *J. Roy. Statist. Soc. Ser. B* **48**, 1-38.

Durairajan, T.M. and B.K. Kale (1979). Locally most powerful test for the mixing proportion. *Sankhyā Ser. B* **41**, 91-100.

Efron, B. (1975). Defining the curvature of a statistical problem. *Ann. Statist.* **3**, 1189-1242.

Ghosh, J.K. and P.K. Sen (1985). On the asymptotic performance of the log likelihood ratio statistic for the mixture model and related results. In: L.M. LeCam and R.A. Olshen, Eds., *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer, II.* Wadsworth, Belmont, CA, 789-806.

Gokhale, D.V. and A. SenGupta (1986). Optimal tests for the correlation coefficient in a symmetric multivariate normal population. *J. Statist. Plann. Inference* **14**, 263-268.

Hartigan, J.A. (1985). A failure of likelihood asymptotics for normal mixtures. In: L.M. LeCam and R.A. Olshen, Eds., *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer, II.* Wadsworth, Belmont, CA, 807–810.

Newcomb, S. (1886). A generalized theory of the combination of observations so as to obtain the best result. *Amer. J. Math.* **8**, 343–366.

Rao, C.R. (1973). *Linear Statistical Inference and its Applications.* Wiley, New York, 2nd ed.

Self, S.G. and K. Liang (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J. Amer. Statist. Assoc.* **82**, 605–610.

SenGupta, A. (1987). On tests for equicorrelation coefficient of a standard symmetric multivariate normal distribution. *Austral. J. Statist.* **29**, 49–59.

SenGupta, A. and L. Vermeire (1986). Locally optimal tests for multiparameter hypotheses. *J. Amer. Statist. Assoc.* **81**, 819–825.

Spjøtvoll, E. (1968). Most powerful test for some non-exponential families. *Ann. Math. Statist.* **39**, 772–784.

Titterington, D.M., A.F.M. Smith and U.E. Makov (1985). *Statistical Analysis of Finite Mixture Distributions.* Wiley, New York.