Lifetime Data Analysis, 6, 281–290, 2000

Optimal Tests for No Contamination in Reliability Models

CHANDRANATH PAL
Department of Statistics, University of Kalyani, West Bengal 741 235, INDIA

ASHIS SENGUPTA ashis@www.isical.ac.in
Applied Statistics Division, Indian Statistical Institute, 203 Barrackpore Trunk Road, Calcutta 700 035, INDIA

Received October 5, 1998; Revised October 25, 1999; Accepted January 31, 2000

Abstract. Inferences on mixtures of probability distributions, in general, and of life distributions, in particular, are receiving considerable importance in recent years. The likelihood ratio procedure of testing for the null hypothesis of no contamination is often very cumbersome and lacks its usual asymptotic properties. Recently, SenGupta (1991) has introduced the notion of an 'L-optimal' test for such testing problems. The idea is to recast the original several parametric hypotheses representation of the null hypothesis in terms of only a single hypothesis involving an appropriately chosen parametric function. This approach is shown to be both mathematically elegant and operationally simple for a quite general class of mixture distributions which contains, in particular, all mixtures of the one-parameter exponential family and also a very rich subclass of mixtures useful in life-testing and reliability analysis. It is also illustrated through two examples—one based on real-life data and the other on a simulated sample.

Keywords: contamination model, optimal tests, Pivotal Parametric Product, reliability distributions

1. Introduction

Mixture distributions are recently playing very important roles in theoretical as well as in applied statistics. While a lot of work has so far been done on the estimation aspect of mixture distributions, the problem of finding optimal tests is only recently being attacked. Mixtures of distributions are also 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 which establish 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, in batches, is acquired 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. Ad hoc procedures for testing for 'no contamination' are subjective and often lack optimality properties. One then has to develop an 'objective' procedure for this purpose which at the same time should preferably be such that in practice it can be applied and interpreted easily. The likelihood ratio procedure is often very cumbersome and the usual asymptotic results for the distribution of the test criterion do not hold. Under a general setup, 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. Recently SenGupta (1991) has introduced the notion of a parametric function called the *Pivotal Parametric Product* (P^3, in short) to characterize the hypothesis of no mixture when both the parameter and the mixing proportion (p) are unknown. Not only simple tests for no contamination can then be constructed but also the distributions of the test statistics, in many cases, can be conveniently derived. Further this test has the property of being L-optimal, i.e., *its power matches that of the locally most powerful (LMP) test for the parameter for each given value of* p (> 0).

The reason for considering 'local' optimality is that there does not exist a test which is 'globally' optimal for all alternative hypothetical values of the parameters. 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, in particular, LMP tests and/or their extensions (*e.g.* $C(\alpha)$-tests) under the presence of nuisance parameters can be conveniently derived and are very much useful (see Kocherlakota and Kocherlakota, 1985). The L-optimal tests are 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.

In Section 2 we give a set of sufficient conditions on the class of densities such that the associated class of mixtures admits an appropriate P^3, 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 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 4 we 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 of Section 2. We also illustrate this result through several distributions useful in reliability theory. For each distribution, we see 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. Finally, in Section 5, we illustrate the application of our procedure through two examples—one on a real-life data set (Barnett and Lewis, 1994) on crushed rock used for road and rail bases and the other on a simulated sample drawn from a mixture of two lognormal distributions.

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

2. The P^3 Approach

Consider the mixture model with density

$$g(x \mid p, \theta) = pf(x \mid \theta) + (1-p)f(x \mid \theta_0) \tag{2.1}$$

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 'contaminating' density $f(x \mid \theta)$ 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: $[H_{01}\colon\ p = 0 \bigcup H_{02}\colon\ \theta = \theta_0 \bigcup H_{03}\colon\ p = 0 \text{ and } \theta = \theta_0]$.

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* (see SenGupta, 1991, for further motivations of this approach). 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 simulation.

Denote by $\mathcal{G}$ the class of density functions g, given by (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 $X_1, \ldots, X_n$ be n *iid* observations drawn from a population with density g.

LEMMA 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 $\partial f(x \mid \theta)/\partial\theta|_{\theta_0}$ exists and is finite for all $x \in \mathcal{X}$.*

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

(C5) $E_{f_\theta}\left(\left.\frac{\partial \log f(X \mid \theta)}{\partial\theta}\right|_{\theta_0}\right) \stackrel{\text{def}}{=} \gamma(\theta, \theta_0) = 0$ *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 $\frac{1}{n}\sum_{i=1}^{n} \frac{\partial \log f(X_i|\theta)}{\partial\theta}|_{\theta=\theta_0}$.

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

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

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 $n \to \infty$. T is therefore consistent for η. ■

COROLLARY 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*

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

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 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. ■

Remark 2.1. (C5) of Lemma 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 $f(x \mid \theta)$ be the density of a von Mises distribution given by

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

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

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

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

3. Integrated Likelihood Approach and the 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

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

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

$$\begin{aligned}\left.\frac{d \log \tilde{L}(\theta)}{d\theta}\right|_{\theta_0} &= \left[\frac{1}{\tilde{L}(\theta)} \frac{d\tilde{L}(\theta)}{d\theta}\right]_{\theta_0} \\ &= \left[\frac{1}{\tilde{L}(\theta)} \frac{d}{d\theta} \int_0^1 \prod_{i=1}^{n} \{p f(x_i \mid \theta) + (1-p) f(x_i \mid \theta_0)\}\pi(p)dp\right]_{\theta_0} \qquad (3.1)\end{aligned}$$

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 (3.1) (see Theorem 7.40 of Apostol, 1974, p. 167). We then have, after some simplification,

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

where

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

Summarizing the above, we have the following

THEOREM 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.*

4. Optimal P^3 Tests in Reliability Models

In view of Lemma 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 4.1 *Consider a class* $\mathcal{F}_0$ *of reliability distributions satisfying conditions (C1)–(C5) of Lemma 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 2.1. Consistency of the test follows from the consistency of T as an estimator of η. ■

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

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

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

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

(iv) Inverse Gaussian with location:
$f^{(4)}(x \mid \theta) = (2\pi x^3)^{-1/2} \exp\{-(x-\theta)^2/2\theta^2 x\};\ \theta_0 = 1.$

(v) Inverse Gaussian with dispersion:
$f^{(5)}(x \mid \theta) = (\theta/2\pi x^3)^{1/2} \exp\{-(x-\theta)^2/2\theta^2 x\};\ \theta_0 = 1.$

(vi) Bivariate Exponential Conditional (BEC):
$f^{(6)}(x, y \mid \theta) = \beta\lambda v(\theta) \exp(\beta x + \lambda y + \theta\beta\lambda xy);\ \theta_0 = 0,$
β, λ are known (each assumed to be equal to 1) (see Arnold and Strauss, 1988).

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

(viii) Bivariate Inverse Gaussian:

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

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

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

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

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 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 $g^{(i)}$, $i = 1, \ldots, 7$:

$$\eta^{(1)} = p\theta;\quad T^{(1)} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

$$\eta^{(2)} = p\theta;\quad T^{(2)} = \frac{1}{n}\sum_{i=1}^{n} \log X_i,$$

$$\eta^{(3)} = p(\theta-1);\quad T^{(3)} = \frac{1}{n}\sum_{i=1}^{n}(X_i - 1),$$

$$\eta^{(4)} = p(\theta-1);\quad T^{(4)} = \frac{1}{n}\sum_{i=1}^{n}(X_i - 1),$$

$$\eta^{(5)} = \frac{p}{2}\left(1-\frac{1}{\theta}\right);\quad T^{(5)} = \frac{1}{2n}\sum_{i=1}^{n}\left(3 - X_i - \frac{1}{X_i}\right).$$

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

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

$f^{(8)}$ through $f^{(10)}$, though not members of the exponential family, may be seen to satisfy the conditions of Lemma 2.1. For $g^{(8)}$, let

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

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,

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

For $g^{(9)}$, since the score function vanishes identically at $\theta_0 = 0$, we use a reparametrization $\theta' = \theta^2$ (SenGupta and Pal, 1993) and note after a little algebra that

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

where

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

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

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

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.

5. 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.

5.1. 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 $\chi^2_{.01,24} = 42.98$ 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.

5.2. 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 ‘mixing’ purpose have been generated by the RNUN subroutine.

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.

Acknowledgment

The authors are thankful to a referee and the associate editor for constructive suggestions, including the one which led to the modified Lemma 2.1 and Theorem 4.1, resulting in this improved presentation of the paper.

References

M. Aitkin and D. B. Rubin, "Estimation and hypothesis testing in finite mixture models," *J. Roy. Statist. Soc. Ser. B* vol. 47 pp. 67–75, 1985.

T. M. Apostol, *Mathematical Analysis, 2nd ed.*, Addison-Wesley Publishing Company: Philippines, 1974.

B. C. Arnold and D. Strauss, "Bivariate distributions with exponential conditionals," *J. Amer. Statist. Assoc.* vol. 83 pp. 522–527, 1988.

V. Barnett and T. Lewis, *Outliers in Statistical Data, 3rd ed.*, John Wiley and Sons, 1994.

H. Block and H. Joe, "Tail behavior of the failure rate functions of mixtures," *Lifetime Data Analysis* vol. 3 pp. 269–288, 1997.

B. S. Everitt, "Mixture distributions," in N. L. Johnson and S. Kotz (eds.), *Encyclopedia of Statistical Sciences*, John Wiley, vol. 5, pp. 559–569, 1985.

R. F. Gunst and R. J. T. Webster, "Density functions of the bivariate chisquare distribution," *J. Statist. Comput. Simul.* vol. 2 pp. 275–288, 1973.

R. Jiang and D. N. P. Murthy, "Modelling failure data by mixture of 2 Weibull distributions: A graphical approach," *IEEE Transactions on Reliability* vol. 44 pp. 477–488, 1995.

N. L. Johnson and S. Kotz, *Continuous Univariate Distributions – 1*, John Wiley and Sons, 1970.

K. Kocherlakota and S. Kocherlakota, "On some tests for independence in nonnormal situations: Neyman's $C(\alpha)$ test," *Commun. Statist. – Theor. Meth.* vol. 14 pp. 1453–1470, 1985.

S. Kocherlakota, "The bivariate inverse Gaussian distribution: An introduction," *Commun. Statist. - Theor. Meth.* vol. 15 pp. 1081–1112, 1986.

E. L. Lehmann, *Theory of Point Estimation*, John Wiley and Sons, 1983.

K. V. Mardia, *Statistics of Directional Data*, Academic Press, 1972.

S. P. Mukherjee and A. Chatterjee, "Estimation of parameters in finite mixtures of exponential families from censored data," *Commun. Statist., Ser. A* vol. 16 pp. 3133–3147, 1987.

B. G. Quinn, G. J. McLachlan and N. L. Hjort, "A note on the Aitkin-Rubin approach to hypothesis testing in mixture models," *J. Roy. Statist. Soc. B* vol. 49 pp. 311–314, 1987.

A. SenGupta, "A review of optimality of multivariate tests," in S. R. Jammalamadaka and A. SenGupta (guest eds.), *Special Issue: Multivariate Optimality and Related Topics, Statistics & Probability Letters*, vol. 12 pp. 527–535, 1991.

A. SenGupta and C. Pal, "Optimal tests in some applied mixture models with non-regularity problems," in S. K. Basu and B. K. Sinha (eds.), *Probability and Statistics*, Narosa Publishing House: New Delhi, pp. 151–164, 1993.

D. M. Titterington, A. F. M. Smith and U. E. Makov, *Statistical Analysis of Finite Mixture Distributions*, John Wiley and Sons, 1985.