Department for Applied Statistics
Johannes Kepler University Linz

IFAS Research Paper Series
2009-39

Exact likelihood ratio testing for homogeneity of the exponential distribution

Milan Stehlík and Helga Wagner

January 2009

Abstract

The aim of this paper is to provide guidelines for homogeneity testing of the exponential distribution against the general (subpopulation model) or more component subpopulation alternatives. We introduce the exact likelihood ratio test of homogeneity in the subpopulation model, ELRH, and the exact likelihood ratio test of homogeneity against the 2-component sample mixture, ELR2. The ELRH is asymptotically optimal in the Bahadur sense when the alternative consists of sampling from a fixed number of components and thus ELRH is in some setups superior to frequently used tests based on EM algorithm like the modified likelihood ratio test, the ADDS test and the D-test among others. We demonstrate this fact by both theoretical comparisons and simulations. A real data example illustrates the methods discussed.

Keywords: Homogeneity testing, exponential distribution, likelihood ratio, subpopulation model, exact distribution, asymptotic efficiency, exact slopes, mixture models

1 Introduction

The exponential distribution is one of the most widely used lifetime distributions in reliability engineering. It has a density of the form

$$f(y_i|\theta_i) = \theta_i \exp(-\theta_i y_i), y_i > 0,$$

where $1/\theta_i > 0$ is a scale parameter of exponential distribution. There is a big body of literature on the theory and applications of the exponential distribution (see Balakrishnan and Basu, 1996). The problem of testing for heterogeneity or overdispersion has received more attention than tests of the number of components (see Susko, 2003). The assumption that the data are generated by a mixture of exponential distributions is widely used in the analysis of lifetime data. The hazard rate of a one-component exponential distribution is constant, whereas the hazard rate of a mixture of exponentials decreases. Therefore the mixture model is frequently adopted to fit the distribution of a time to failure where the observed failure rate seems to decline with time. Often the mixture can be explained by competing risks. The components in the mixture correspond to the distinct causes of failure which are taken to act in a mutually exclusive manner. For example, Choi (1979) used a two-component mixture model to study the toxicity of chemical agents. For a survey of mixtures of exponentials see McLachlan (1995).

The aim of this paper is to introduce the efficient procedure for testing exponential homogeneity against alternatives of exponential heterogeneity. The likelihood-ratio decision procedure related to the hypothesis $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1 \setminus \Theta_0$, $\emptyset \neq \Theta_0 \subset \Theta_1 \subset \Theta$, is based on the ratio

$$\frac{\sup_{\theta\in\Theta_0} L_y(\theta)}{\sup_{\theta\in\Theta_1} L_y(\theta)}$$

where $\Theta_0 \subset \Theta_1$, θ is the parameter of interest and $L_y(\theta)$ is the likelihood of θ under the observed data y. Alternatives to homogeneity are often specified as mixture models. The likelihood of a sample $y_1, \ldots, y_N$ of iid observations from a 2 component mixture is given as

$$f(y_1, \ldots, y_N) = \prod_{i=1}^{N} \left[p\theta_1 \exp(-\theta_1 y_i) + (1-p)\theta_2 \exp(-\theta_2 y_i) \right], 0 < p < 1$$

and the likelihood of a sample from a general $k - component$ mixture of exponential components is

$$f(y_1, \ldots, y_N) = \prod_{i=1}^{N} \left(\sum_{j=1}^{k} p_j \theta_j \exp(-\theta_j y_i) \right),$$

where $0 < p_j < 1, \sum_j p_j = 1$. For the mixture alternative, there exist a very concrete justification for using the likelihood ratio tests: the likelihood ratio test is consistent against all alternatives with decreasing failure rate (see Randles, 1982; Tchirina, 2005).

In this paper, we consider two exact likelihood ratio tests, the exact likelihood ratio test of homogeneity (ELRH) for the general subpopulation model proposed by Stehlík (2003) and the exact likelihood ratio test for 2 subpopulations (ELR2), proposed by Stehlík and Ososkov (2003). Generalization to testing for k subpopulations, ELRk, is theoretically easy to implement, however it is computationally expensive and computational difficulty increases with k. However, some applications of ELRk can be found in physics, see e.g. Efimova et al. (1989).

In the subpopulation model the number of subpopulations has to be specified, so the "N subpopulation" model would be defined by the joint density

$$f(y_1, \ldots, y_N) = \prod_{i=1}^{N} \theta_i \exp(-\theta_i y_i) \tag{1}$$

which is the alternative tested in the ELRH test. The ELR2 test uses the alternative of two subpopulations, which can be specified by (1) and two nonempty index sets M_1, M_2 such that

$$M_1 \cup M_2 = \{1, ..., N\}, M_1 \cap M_2 = \emptyset \tag{2}$$

$$\forall j \in M_1 : \theta_j = \theta_1, \forall j \in M_2 : \theta_j = \theta_2, \theta_1 \neq \theta_2. \tag{3}$$

Note that the alternative in the subpopulation model is exponential heterogeneity in the sample, whereas the mixture alternative can be interpreted as exponential heterogeneity in the population.

The reason why we consider the subpopulation model is, besides simplicity, the fact that as soon the difference between the number of components in the mixture model under H_0 and H_1 respectively is greater than 1, the likelihood ratio tests involves nonstationary random fields, for which very few theoretical results are available (see Garel, 2007). The ELRH-test is asymptotically optimal in the Bahadur sense when the alternative consists of the subpopulation model with a finite number of populations (see Stehlík, 2006; Rublík, 1989a,b). ELRH and ELR2 tests have nonstandard asymptotic distributions but their exact distribution can easily be simulated. Notice, that our setup encompasses also the case of the Weibull distribution with known shape parameter. The exact tests of homogeneity when the shape parameter of the Weibull is unknown are not known to the authors. Tests for exponentiality against Weibull alternative are given in Meintanis (2007) and Henze and Meintanis (2005).

The paper is organized as follows. In section 2 the exact likelihood ratio homogeneity tests ELRH and ELR2 are introduced and discussed. In section 3 a comparative power study of tests of homogeneity is provided together with the theoretical explanation of the obtained results. Here other tests for homogeneity in exponential mixtures are dispersion score (DS) tests (also known under the name $C(\alpha)$-tests; see chapter 4 of Lindsay (1995)) and recently proposed modified likelihood ratio test (MLRT) introduced by Chen et al. (2001), which is a penalized LRT and has standard χ^2 asymptotics, the ADDS test by Mosler and Seidel (2001), a combination of the dispersion score test with a properly chosen goodness-of-fit procedure and the D-test by Charnigo and Sun (2004), based on the L^2 distance between the estimated densities of a homogeneous and a heterogeneous model. Charnigo and Sun (2004) also introduce a penalized and several weighted variants of the D-test. An illustrative example follows in Section 4. To maintain the continuity of explanation the technicalities and proofs are put into Appendix.

2 Homogeneity testing

2.1 The ELRH test

For a sample of N independent observations $y = (y_1, \ldots, y_N)$, where $y_i \sim$ Exponential (θ_i) we consider the LR homogeneity test against the general alternative (subpopulation model), i.e.

$$H_0 : \theta_1 = \ldots = \theta_N \text{ versus non} H_0.$$

The exact distribution of the LR test of homogeneity against the general alternative, the ELRH test, was derived in Stehlík (2006) for the exponential and Weibull distribution and for the generalized gamma distribution in Stehlík (2008). LR tests have good properties (see e.g. Lehmann, 1964; Manoukian, 1986) and in regular cases they are optimal (see also Appendix). The LR statistics $-\ln \Lambda_N$ is derived in Theorem 3 of Stehlík (2006) for $y_1, \ldots, y_N$ i.i.d. from the exponential distribution. It has the form

$$N \ln(\sum_{i=1}^{N} y_i) - N \ln N - \sum_{i=1}^{N} \ln y_i.$$

A very important property of the LR test of homogeneity is its scale invariance, i.e. its distribution under H_0 is independent of the unknown scale parameter. This is an advantage in comparison to some asymptotical tests and tests depending on the true but unknown value of θ. The critical values are easy to obtain by simulation, e.g. from the standard exponential distribution or the Dirichlet distribution. Table 1 gives the critical values for $N = 20, 50, 100, 500$ obtained by simulation. $M = 1000000$ samples of size N were generated yielding a sample of $c_1, \ldots, c_M$ for the test statistic $-\ln \Lambda_N$ under homogeneity. The values $c_{1-\alpha}$ are determined as the respective order statistic $c_{1-\alpha} = c_{\left(M(1-\alpha)\right)}$. These values are used throughout the paper.

Table 1: Critical values for the ELRH

N	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$
10	8.5279	9.7994	12.4814
20	15.7276	17.3881	20.7894
50	35.7658	38.2121	43.0112
100	67.6812	70.9487	77.4036
500	311.3698	318.2312	331.4319
1000	609.4173	619.1025	637.3055

The log LR under homogeneity is a monotonous function of the statistic $\frac{N^N y_1 \dots y_N}{(y_1+\dots+y_N)^N}$ which is clearly scale independent. The LR homogeneity statistic is also a monotonous transformation of the so called Moran's statistics T_N^+ and T_N^-, where

$$T_N^+ = C + \frac{1}{N}\sum_{i=1}^{N} \ln \frac{x_i}{\bar{x}},$$

$T_N^- = -T_N^+$, $C \approx 0.57721566$ is the Euler constant and $\bar{x}$ is the arithmetic mean. This is a scale-free exponentiality test for which large deviations are studied in Tchirina (2005). The test of homogeneity provided in this paper are asymptotically optimal in the Bahadur sense (see Rublík, 1989a,b; Stehlík, 2003) when the underlying distribution is exponential and when the alternative of the homogeneity consists of sampling from a fixed number m of populations with relative sample sizes n_j/N tending to positive limits p_j (subpopulation model).

2.1.1 Simulation Study

A simulation study was conducted to determine the power of the test for a mixture of two exponential components with pdf

$$f(y) = p\exp(-y) + (1-p)\theta\exp(-\theta y)$$

for $\theta = 1, 2, \dots, 10$ and different component weights $p = 0.1, 0.5, 0.9$. We used two different sizes of the test, namely $\alpha = 0.01$ and $\alpha = 0.05$ and $N = 20, 50, 100$. For each parameter combination $M = 10000$ samples were generated and the proportion of rejections of the ERLH test was determined.

Results given in Table 2 show that the ELRH test holds the chosen size α also for small samples. The power of the ELRH test increases with θ, for fixed θ the highest power is obtained for equal component weights, whereas for $p = 0.9$ the power can be rather low, in particular considerably lower than for $p = 0.1$. This behaviour of the power is not specific to the ELRH test but has been noted for different homogeneity tests in Mosler and Haferkamp (2007). It can be explained by interpreting the mixture as a contaminated distribution: if $p = 0.1$ the density of the second component with parameter θ is predominant. The mixture with component 1 leads to a modification of the Exponential (θ) in the tail region. If however $p_1 = 0.9$, the first component is predominant, and mixing has an impact on the density close to the mode which is 0. Mosler and Haferkamp (2007) refer to the first case as 'upper' and to the second case as 'lower' contamination. Lower contamination is hard to detect as the mixture distribution is hardly different from the (homogeneous) exponential distribution.

2.2 The ELR2 test

In this section we will discuss the efficient testing procedure of the number of components m in the Exponential mixture for $m = 2$ firstly introduced by Stehlík and Ososkov (2003). Testing for the number of components involves inference for an overfitting mixture model and in this case represents a nonregular problem (see Frühwirth-Schnatter, 2006, Section 4.2).

Here we consider the LR homogeneity testing with more a complex alternative H_1, which is the approximation to a finite scale mixture. In physics, such testing corresponds to the testing of the number m of secondary particles obtained after the collision (under the condition that the reasonable conservation of energy had been present during the collision). Then homogeneity corresponds to one particle ($m = 1$), and $m > 1$ corresponds to m particles (cf Efimova et al., 1989; Stehlík and Ososkov, 2003). We consider the alternative of the form $H_1 : m = 2$. The hypothesis

$$H_0 : m = 1 \text{ versus } H_1 : m = 2 \tag{4}$$

in the mixture model can be approximated (following Stehlík and Ososkov (2003)) by the hypothesis of the subpopulation model

$$H_0 : \ \theta_1 = ... = \theta_n \text{ versus } H_1 : \exists M_1, M_2, M_1 \cup M_2 = \{1, ..., N\}, \tag{5}$$

$$M_1 \cap M_2 = \emptyset, M_1, M_2 \neq \emptyset, \forall j \in M_1 : \theta_j = \theta_1, \forall j \in M_2 : \theta_j = \theta_2, \theta_1 \neq \theta_2$$

e.g. by the null hypothesis of the homogeneity with the modified alternative, which is actually a subset of the alternative of the hypothesis of the homogeneity. We construct the LR test of the hypothesis (5) which approximates the hypothesis (4).

2.2.1 The exact LR test

Let $y_1, ..., y_N$ be independently distributed with exponential densities. The LR of the test of the hypothesis (5) has the form

$$\Lambda_N(y) = \frac{\max_{\theta_1 = ... = \theta_N} f(y, \theta)}{\max_{H_1} f(y, \theta)},$$

where $y = (y_1, ..., y_N), \theta = (\theta_1, .., \theta_N)$. To compute the denominator $\max_{H_1} f(y, \theta)$ we proceed as follows. Suppose that $\{y_{i_1}, .., y_{i_K}\}$, $0 < K < N$ are the observations from the exponential distribution with the scale parameter θ_1 and the other observations are distributed according to the exponential distribution with the scale parameter θ_2. Following Stehlík and Ososkov (2003) we obtain the formula

$$\Lambda_N(y) = \min_{0<K<N, p \in P(K)} \{ \frac{N^N}{K^K (N-K)^{N-K}} \frac{(y_{i_1} + ... + y_{i_K})^K (y_{i_{K+1}} + ... + y_{i_N})^{N-K}}{(y_1 + ... + y_N)^N} \}, \tag{6}$$

Table 2: Simulated power for the ELRH-test

θ	1	2	3	4	5	6	7	8	9	10
	$\alpha = 0.05$									
	$p = 0.1$									
N=20	0.0533	0.0681	0.1166	0.1628	0.2312	0.2882	0.3287	0.3796	0.4302	0.4679
N=50	0.0481	0.0703	0.1612	0.2878	0.4142	0.5036	0.6050	0.6509	0.7260	0.7556
N=100	0.0580	0.0989	0.2349	0.4226	0.5770	0.6954	0.8064	0.8641	0.9080	0.9299
N=1000	0.0565	0.2859	0.8403	0.9930	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	$p = 0.5$									
N=20	0.0620	0.0841	0.1914	0.2922	0.3823	0.4861	0.5845	0.6632	0.7430	0.7843
N=50	0.0471	0.1141	0.3188	0.5089	0.7146	0.8289	0.8974	0.9522	0.9752	0.9844
N=100	0.0517	0.1742	0.5130	0.7717	0.9338	0.9781	0.9982	0.9970	1.0000	1.0000
N=1000	0.0568	0.7281	0.9996	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	$p = 0.9$									
N=20	0.0527	0.0625	0.0819	0.0875	0.1089	0.1215	0.1353	0.1598	0.1583	0.1805
N=50	0.0484	0.0621	0.0964	0.1201	0.1554	0.1847	0.2309	0.2584	0.2745	0.3085
N=100	0.0461	0.0902	0.1206	0.1734	0.2472	0.2856	0.3467	0.3887	0.4369	0.4611
N=1000	0.0545	0.1800	0.4803	0.7556	0.9082	0.9703	0.9907	0.9970	0.9991	0.9998
	$\alpha = 0.01$									
	$p = 0.1$									
N=20	0.0098	0.0161	0.0324	0.0597	0.0893	0.1501	0.1869	0.2278	0.2615	0.3111
N=50	0.0079	0.0184	0.0468	0.1221	0.2235	0.3110	0.4164	0.5109	0.5645	0.6447
N=100	0.0120	0.0305	0.0877	0.2212	0.3811	0.5342	0.6431	0.7461	0.8074	0.8756
N=1000	0.0101	0.1070	0.6498	0.9718	0.9994	1.0000	1.0000	1.0000	1.0000	1.0000
	$p = 0.5$									
N=20	0.0098	0.0244	0.0595	0.1218	0.1755	0.2492	0.3277	0.4023	0.4660	0.5278
N=50	0.0117	0.0348	0.1164	0.2785	0.4628	0.6320	0.7386	0.8438	0.8867	0.9341
N=100	0.0124	0.0672	0.2631	0.5558	0.7881	0.9114	0.9710	0.9911	0.9964	0.9994
N=1000	0.0127	0.4710	0.9994	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	$p = 0.9$									
N=20	0.0105	0.0162	0.0211	0.0190	0.0401	0.0388	0.0423	0.0563	0.0508	0.0621
N=50	0.0092	0.0156	0.0236	0.0346	0.0403	0.0639	0.0663	0.0941	0.1018	0.1148
N=100	0.0096	0.0233	0.0388	0.0564	0.0842	0.1080	0.1383	0.1642	0.2066	0.2359
N=1000	0.0118	0.0553	0.2205	0.5011	0.7221	0.8726	0.9423	0.9734	0.9893	0.9958

where $P(K)$ denotes all partitions of $\{1,..,K\}$ in two nonempty subsets. The main advantages of the test statistic (6) is that under the H_0 it does not depend on the unknown value of the parameter θ. The distribution of the LR test statistics $-\ln \Lambda_N$ where Λ_N is given by the formula (6) under the null hypothesis is derived in Stehlík and Ososkov (2003). The main advantage of this representation is that the distribution of the LR statistics $-\ln \Lambda_N$ can be obtained very easily by simulation requiring only draws from a standard exponential distribution. For the alternative of the two-component mixture form, i.e. for the testing problem

$$H_0 : \theta_0 \exp(-\theta_0 x) \text{ versus } H_A : p\theta_0 \exp(-\theta_0 x) + (1-p)\theta \exp(-\theta x), \theta > \theta_0, 0 < p \le 1, \quad (7)$$

where θ_0 is known and θ, p are unknown parameters we will get the similar asymptotical behavior as was derived by Hartigan (1985). He discovered the divergence of the LR test statistics for testing homogeneity in normal mean mixture models with an unbounded mean parameter. Liu et al. (2003) have proved in the setup (7) that $\lim_{N\to\infty} P(2\Lambda_N - \log\log N + \log(16\pi^2) \le x) = \exp(-\exp(-x/2))$. They also try to determine whether it is feasible to approximate $2\Lambda_N - \log\log N + \log(16\pi^2)$ by the extreme value distribution for a large N. Unfortunately, as they reported in Liu et al. (2003) this approximation is quite poor even for a sample size as large as 5000. Therefore we suggest to use ELR2 (subpopulation model) with critical values simulated from the exact distribution. The following subsection provides some notes on determination of ELR2 test statistics.

2.2.2 Determination of the Likelihood ratio statistic

The determination of the likelihood-ratio statistic of the ELR2 test

$$\begin{aligned} -\ln \Lambda_N(y) = \; & -\min\nolimits_{0<K<N, p\in P(K)}\{N\ln N - K\ln K - (N-K)\ln(N-K) + \\ & +K\ln(\textstyle\sum_{n=1}^{K} y_{i_n}) + (N-K)\ln(\sum_{n=1}^{N-K} y_{i_n}) - N\ln(\sum_{n=1}^{N} y_n)\} \end{aligned}$$

is not so straight forward as for the homogeneity test of section 2, as the minimum of

$$N\ln N - K\ln K - (N-K)\ln(N-K) + K\ln(\sum_{n=1}^{K} y_{i_n}) + (N-K)\ln(\sum_{n=1}^{N-K} y_{i_n}) - N\ln(\sum_{n=1}^{N} y_n)$$

over all possible classifications into 2 non-empty groups has to be found. For determining this minimum

$$N\ln N - N\ln(\sum_{n=1}^{N} y_n)$$

is irrelevant and therefore the minimum of

$$H(y,K) = \{-K\ln K - (N-K)\ln(N-K) + K\ln(\sum_{n=1}^{K} y_{i_n}) + (N-K)\ln(\sum_{n=1}^{N-K} y_{i_n})\}$$

for $0 < K < N, p \in P(K)$ is of interest. For N observations there are $2^{N-1}-1$ different classifications into 2 nonempty groups. Hence direct minimization over all

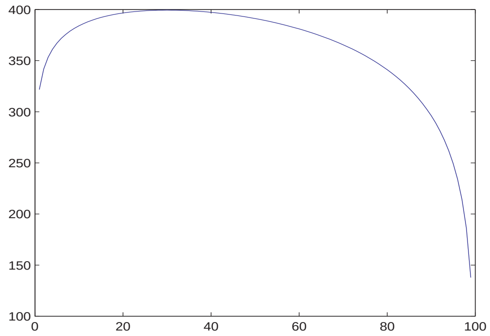

Figure 1: The function $H(x|K)$

classifications is feasible only for small N. However minimizing first $H(y|K)$ for fixed $K = 1, \dots, N-1$ and then determining the minimum of these $N-1$ values poses no problem. Given the data y, their sum $S = \sum_{n=1}^{N} y_i$ is fixed. For fixed K we therefore consider minimization of

$$H(x|K) = K \ln x + (N-K) \ln(S-x)$$

as a function of x. This function is continuous and strictly concave, as the second derivative

$$\frac{\partial^2 H(x|K)}{\partial x^2} = -\Big(\frac{K}{x^2} + \frac{N-K}{(S-x)^2}\Big)$$

is negative for $0 < x < S$. The maximum of $H(x|K)$ is attained at $x = KS/N$ and the minimum lies on one of the boundaries, see Figure 1. For the likelihood ratio statistic, x can only take certain discrete values, as $x = \sum_{n=1}^{K} y_{i_n}$. Miminum and maximum of x are the sum of the smallest and largest K order statistics respectively:

$$\min(x) = \sum_{i=1}^{K} y_{(i)} \qquad \max(x) = \sum_{i=1}^{K} y_{(N-i+1)}$$

Due to the symmetry

$$K \ln(\sum_{i=1}^{K} y_{(N-i+1)}) + (N-K) \ln(\sum_{i=K+1}^{N} y_{(N-i+1)}) = K' \ln(\sum_{i=1}^{K'} y_{(i)}) + (N-K') \ln(\sum_{i=K'+1}^{N} y_{(l)})$$

for $K' = N-K$ the minimum value of $H(x|K)$ is

$$H_{\min} = \min_{0<K<N} (-K \ln K - (N-K) \ln(N-K) + K \ln(\sum_{i=1}^{K} y_{(i)}) + (N-K) \ln(\sum_{i=K+1}^{N} y_{(i)})$$

which can be determined very simply as only sums of order statistics are involved. The ELR2 test statistic is given as

$$\ln \Lambda_N(y) = N \ln N - N \ln(\sum_{n=1}^{N} y_n) + H.$$

Critical values can be obtained similarly to the homogeneity test by generating M samples of size N from the standard exponential distribution, computing the test statistic for each sample and determining $c_{1-\alpha}$ by the respective order statistic $c_{1-\alpha} = c_{(M(1-\alpha))}$.

3 Comparative power study of homogeneity tests

3.1 Simulation Setup

A simulation study with a similar setup as in Mosler and Haferkamp (2007) was performed to compare the power of the exact likelihood ratio tests ELRH and ELR2 to other different tests for the hypotheses

$$H_0 : y_1, \ldots, y_N \sim \text{Exponential}\,(\theta) \tag{8}$$

against the alternative

$$H_1 : y_1, \ldots, y_N \text{ follow a mixture of two exponential components} \tag{9}$$

To be consistent with the simulation setup in Mosler and Haferkamp (2007) we generate proportions $N_1, N - N_1$ in such a way that N_1 has a binomial distribution $Binomial(N, p)$. We used three typical mixture proportions, $p = 0.1, 0.5, 0.9$ and two sample sizes $N = 100$ and $N = 1000$. The parameter of the first mixture component is $\theta_1 = 1$ and for θ_2 a range of values greater than θ_1 was chosen. These settings correspond to upper contamination ($p = 0.1$), fifty-fifty mixtures and lower contamination ($p = 0.9$). For each parameter setting $M = 10000$ samples of N observations from the mixture distribution with density

$$f(y) = p \exp(-y) + (1 - p)\theta \exp(-\theta y)$$

were generated. We compared the proportion of rejections of the null hypothesis for the following tests: the modified likelihood ratio test of Chen et al. (2001) (MLRT); the D-test introduced by Charnigo and Sun (2004) (DTEST) and 2 weighted variants of the D-test (W1D and W2D); the Anderson Darling test (AD), the Dispersion Score test (DST), the combination of AD and DS-test introduced by Mosler and Haferkamp (2007) (ADDS) and the exact likelihood ratio tests against the general alternative (ELRH) and for testing against the 2 component subpopulation model (ELR2).

We also made comparisons to penalized D-test (PenD), but PenD does not hold the size for $N < 1000$ and then it starts to hold it only approximately. Therefore we decided to delete it from our comparisons. Furthermore, as was personally communicated to us by Richard Charnigo, the PenD test is rather anticonservative, especially when looking at data from other (i.e., non-exponential) distributions. However this phenomenon dissipates with larger sample sizes. For small N the same problem occurs for the D-test, because of estimating the nuisance scale parameter. Here we can take an advantage of scale invariance of pivotal statistics based ELRH and ELR2 tests.

Critical values for the AD and DS were determined by simulation from $M = 50000$ samples as in Mosler and Haferkamp (2007). The ADDS-test uses critical

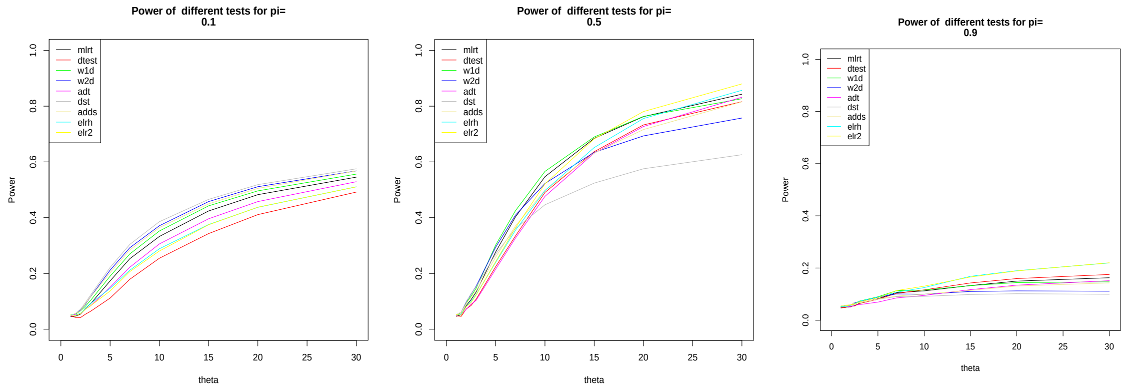

Figure 2: Proportion of rejections of the null hypothesis for different homogeneity tests for $N = 10$, $\alpha = 0.05$ and $p = 0.1$ (left), $p = 0.5$ (middle) and $p = 0.9$ (right)

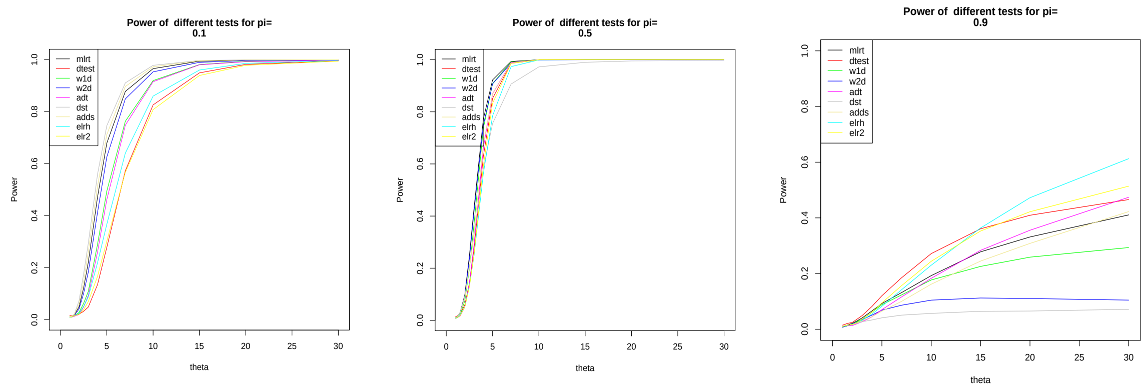

Figure 3: Proportion of rejections of the null hypothesis for different homogeneity tests for $N = 100$, $\alpha = 0.01$ and $p = 0.1$ (left), $p = 0.5$ (middle) and $p = 0.9$ (right)

values for the AD and the DS-test. The interpretation of the behavior of simulated power can be based on exact slopes of the sequence of tests.

3.2 Results

Results of this simulation study for $N = 10$, $N = 100$ and $N = 1000$ are shown in Figures 2–4. Tables 3 and 4 report the simulated powers of the different homogeneity tests ($\alpha = 0.01$) of exponential mixtures together ordered by decreasing powers for $N = 100$ and $N = 1000$, respectively.

For upper contaminations the DST outperforms the other tests for all three sample sizes, followed by W2D- and ADDS-test for $N = 10$ and ADDS and MRLT for $N = 100$ and $N = 1000$. The ERL-tests are among those with lowest power, with ERLH performing better than ELR2. Only the D-test has smaller power than ELRH for $N = 10$ and $N = 1000$.

In fifty-fifty mixtures MLRT performs well for all sample sizes. The weighted variants of the D-tests (W1D and W2D) have relatively high power for large samples ($N = 100$ and $N = 1000$), but only for smaller values of θ for $N = 10$. In small

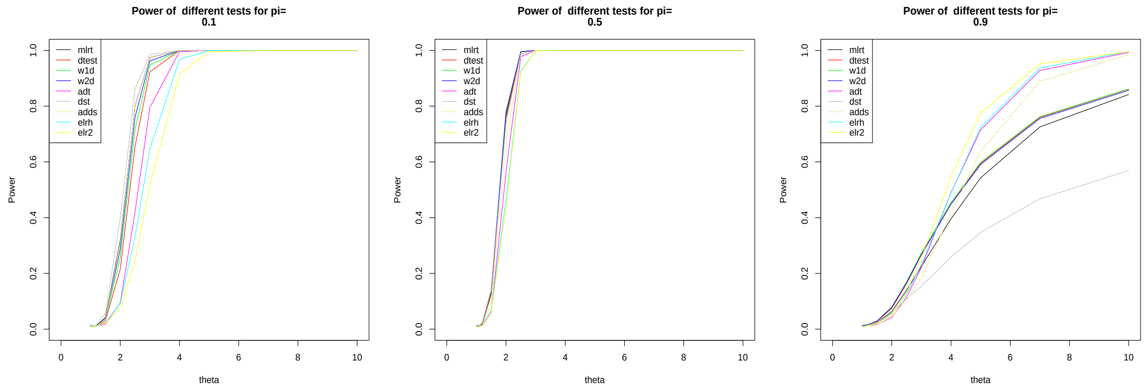

Figure 4: Proportion of rejections of the null hypothesis for different homogeneity tests for $N = 1000$,$\alpha = 0.01$ and $p = 0.1$ (left), $p = 0.5$ (middle) and $p = 0.9$ (right)

samples of mixtures with $\theta > 20$ the ERL2-test is one attaining the highest power, but in larger samples the ERL tests again are among those with lowest power.

All tests have problems to detect heterogeneity in the lower contamination case, in particular for small sample size. In this situation however ERL tests clearly outperform their competitors for larger θ. For $N = 100$ and small θ the D-Test and for $N = 1000$ its weighted variants attain highest power among the compared tests. The dispersion score test is the test with lowest power in a wide range of settings with dramatically lower power than the best test for larger θ.

Particularly we can observe this as a consequence of subexponential model used by the ELR tests and a mixture model by MLRT. All other tests use some kind of 'distance' between the densities which is small for lower contamination.

3.3 Theoretical explanation of the test behavior

Let us have an m-subpopulation model, i.e, the set of overall parameters Θ consists of the m-tuples $\theta = (\theta_1, ..., \theta_m)$, where θ_j is the parameter of the j-th population. Let us suppose that in the k-th experiment the size of sample from the j-th population is $n_k^{(j)}, j = 1, ..., m$ and $k = 1, 2, ...$ Let the product measure $P_{\theta_j}^{\infty}$ corresponds to the (infinite) sampling from the distribution P_{θ_j} (in our case P_{θ_j} has density $\theta_j \exp(-\theta_j y_j), y_j > 0$ with respect to Lebesgue measure). The product measure $P_\theta = P_{\theta_1}^{\infty} \times ... \times P_{\theta_m}^{\infty}$, can be used for description of limiting distribution of independent sampling from these m populations. Thus $n_k = \sum_{j=1}^{m} n_k^{(j)}$ is the total sample size in the k-th experiment. Here we employ the assumption of a finite subpopulation plan given by Rublík (1989b) which together with other regularity conditions guarantees the asymptotical optimality in the Bahadur sense (AOBS) of ELRH test (see Appendix):

i) if $k \neq l$ then $n_k^{(j)} \neq n_l^{(j)}$ for some j

ii) $\lim_{k\to\infty} n_k = +\infty$

iii) $\lim_{k\to\infty} \frac{n_k^{(j)}}{n_k} = p_j \in (0, 1], j = 1, ..., m$

For $\theta = (\theta_1, ..., \theta_m), \theta^\star = (\theta_1^\star, ..., \theta_m^\star) \in \Theta$ let $K(\theta, \theta^\star) = \sum_{j=1}^{m} p_j K(\theta_j, \theta_j^\star)$ where

Table 3: Simulated power for different tests of homogeneity in exponential mixtures
$n = 100$

Rank	Test	Power	Test	Power	Test	Power	Test	Power	Test	Power
					$p = 0.1$					
	$\theta = 1$		$\theta = 2$		$\theta = 3$		$\theta = 5$		$\theta = 10$	
1	dtest	(0.016)	dst	(0.072)	dst	(0.293)	dst	(0.746)	dst	(0.978)
2	adds	(0.012)	adds	(0.058)	adds	(0.260)	adds	(0.711)	adds	(0.973)
3	mlrt	(0.011)	mlrt	(0.049)	mlrt	(0.220)	mlrt	(0.679)	mlrt	(0.965)
4	adt	(0.010)	w2d	(0.041)	w2d	(0.183)	w2d	(0.625)	w2d	(0.953)
5	elr2	(0.010)	w1d	(0.028)	w1d	(0.108)	w1d	(0.494)	w1d	(0.919)
6	dst	(0.010)	elrh	(0.024)	adt	(0.093)	adt	(0.461)	adt	(0.915)
7	elrh	(0.010)	dtest	(0.021)	elrh	(0.087)	elrh	(0.367)	elrh	(0.860)
8	w2d	(0.009)	elr2	(0.020)	elr2	(0.070)	elr2	(0.301)	dtest	(0.826)
9	w1d	(0.009)	adt	(0.020)	dtest	(0.048)	dtest	(0.278)	elr2	(0.808)
					$p = 0.5$					
	$\theta = 1$		$\theta = 2$		$\theta = 3$		$\theta = 5$		$\theta = 7$	
2	dtest	(0.013)	mlrt	(0.098)	mlrt	(0.431)	mlrt	(0.922)	mlrt	(0.993)
3	adds	(0.011)	dst	(0.093)	w2d	(0.417)	w1d	(0.908)	w1d	(0.991)
4	elrh	(0.010)	w2d	(0.093)	w1d	(0.371)	w2d	(0.907)	adt	(0.990)
5	adt	(0.010)	adds	(0.084)	adds	(0.341)	adt	(0.869)	adds	(0.987)
6	mlrt	(0.010)	w1d	(0.074)	dst	(0.328)	adds	(0.869)	w2d	(0.987)
7	dst	(0.009)	dtest	(0.059)	adt	(0.290)	dtest	(0.851)	dtest	(0.987)
8	elr2	(0.009)	adt	(0.051)	dtest	(0.276)	elr2	(0.833)	elr2	(0.986)
9	w1d	(0.008)	elrh	(0.051)	elr2	(0.253)	elrh	(0.785)	elrh	(0.973)
10	w2d	(0.008)	elr2	(0.050)	elrh	(0.248)	dst	(0.755)	dst	(0.907)
					$p = 0.9$					
	$\theta = 1$		$\theta = 2$		$\theta = 5$		$\theta = 10$		$\theta = 30$	
1	dtest	(0.014)	dtest	(0.024)	dtest	(0.120)	dtest	(0.272)	elrh	(0.613)
2	adt	(0.010)	mlrt	(0.021)	elr2	(0.095)	elr2	(0.244)	elr2	(0.514)
3	adds	(0.010)	w2d	(0.019)	mlrt	(0.092)	elrh	(0.230)	adt	(0.474)
4	elr2	(0.010)	elrh	(0.018)	w1d	(0.088)	mlrt	(0.193)	dtest	(0.466)
5	mlrt	(0.010)	dst	(0.018)	elrh	(0.082)	adt	(0.183)	adds	(0.423)
6	dst	(0.010)	adds	(0.018)	adt	(0.070)	w1d	(0.177)	mlrt	(0.411)
7	elrh	(0.010)	elr2	(0.017)	w2d	(0.068)	adds	(0.162)	w1d	(0.293)
8	w1d	(0.008)	w1d	(0.017)	adds	(0.065)	w2d	(0.104)	w2d	(0.105)
9	w2d	(0.007)	adt	(0.013)	dst	(0.041)	dst	(0.057)	dst	(0.072)

Table 4: Simulated power for different tests of homogeneity in exponential mixtures
$n = 1000$

	Test	Power	Test	Power	Test	Power	Test	Power	Test	Power
					$p = 0.1$					
	$\theta = 1$		$\theta = 1.5$		$\theta = 2$		$\theta = 2.5$		$\theta = 3$	
1	w1d	(0.013)	dst	(0.056)	dst	(0.400)	dst	(0.861)	dst	(0.985)
2	w2d	(0.013)	mlrt	(0.041)	mlrt	(0.319)	mlrt	(0.802)	mlrt	(0.974)
3	dtest	(0.013)	w2d	(0.041)	adds	(0.311)	adds	(0.802)	adds	(0.974)
4	dst	(0.011)	w1d	(0.037)	w2d	(0.287)	w2d	(0.761)	w2d	(0.962)
5	adt	(0.011)	adds	(0.035)	w1d	(0.256)	w1d	(0.718)	w1d	(0.947)
6	elr2	(0.010)	dtest	(0.030)	dtest	(0.215)	dtest	(0.656)	dtest	(0.923)
7	mlrt	(0.010)	elrh	(0.022)	adt	(0.094)	adt	(0.421)	adt	(0.796)
8	elrh	(0.010)	elr2	(0.022)	elrh	(0.093)	elrh	(0.331)	elrh	(0.647)
9	adds	(0.009)	adt	(0.019)	elr2	(0.075)	elr2	(0.246)	elr2	(0.517)
					$p = 0.5$					
	$\theta = 1$		$\theta = 1.2$		$\theta = 1.5$		$\theta = 2$		$\theta = 2.5$	
1	adt	(0.013)	dst	(0.022)	dst	(0.141)	w2d	(0.781)	w2d	(0.996)
2	w1d	(0.012)	w1d	(0.022)	w2d	(0.140)	w1d	(0.774)	w1d	(0.995)
3	w2d	(0.012)	w2d	(0.022)	w1d	(0.136)	mlrt	(0.767)	mlrt	(0.995)
4	dtest	(0.011)	dtest	(0.020)	mlrt	(0.131)	dtest	(0.756)	dtest	(0.995)
5	elr2	(0.009)	mlrt	(0.019)	dtest	(0.123)	dst	(0.740)	dst	(0.988)
6	dst	(0.009)	elrh	(0.015)	adds	(0.095)	adds	(0.676)	adds	(0.986)
7	elrh	(0.009)	elr2	(0.014)	elrh	(0.068)	adt	(0.563)	adt	(0.978)
8	mlrt	(0.009)	adds	(0.013)	elr2	(0.061)	elrh	(0.464)	elrh	(0.926)
9	adds	(0.008)	adt	(0.013)	adt	(0.059)	elr2	(0.441)	elr2	(0.923)
					$p = 0.9$					
	$\theta = 1$		$\theta = 2$		$\theta = 3$		$\theta = 5$		$\theta = 10$	
1	w1d	(0.013)	w1d	(0.080)	w1d	(0.274)	elr2	(0.776)	elrh	(0.996)
2	w2d	(0.012)	w2d	(0.079)	w2d	(0.268)	elrh	(0.724)	elr2	(0.995)
3	dtest	(0.012)	dtest	(0.075)	dtest	(0.266)	adt	(0.715)	adt	(0.993)
4	adt	(0.011)	mlrt	(0.063)	elr2	(0.255)	adds	(0.637)	adds	(0.985)
5	elrh	(0.010)	dst	(0.057)	elrh	(0.231)	w1d	(0.599)	w1d	(0.862)
6	elr2	(0.010)	elrh	(0.055)	mlrt	(0.225)	dtest	(0.595)	dtest	(0.861)
7	mlrt	(0.009)	elr2	(0.053)	adt	(0.222)	w2d	(0.591)	w2d	(0.857)
8	dst	(0.008)	adds	(0.046)	adds	(0.186)	mlrt	(0.543)	mlrt	(0.842)
9	adds	(0.007)	adt	(0.042)	dst	(0.154)	dst	(0.347)	dst	(0.570)

the Kullback-Leibler information is defined by the formula

$$K(\theta_j, \theta_j^\star) := \begin{cases} \int \ln \dfrac{dP_{\theta_j}}{dP_{\theta_j^\star}} dP_{\theta_j}, & \text{if } P_{\theta_j} << P_{\theta_j^\star}, \\ +\infty, & \text{otherwise.} \end{cases}$$

Let $\Theta_0 \subset \Theta_1 \subset \Theta$. Then according to the Bahadur-Raghavachari inequality for the exact slope the inequality (see Appendix) $c_T(\theta) \leq 2K(\theta, \Theta_0)$ holds. Here $K(\theta, \Theta_0) := \inf\{K(\theta, \theta^\star) : \theta^\star \in \Theta_0\}$. If $c_T(\theta) = 2K(\theta, \Theta_0)$ for all $\theta \in \Theta_1 \setminus \theta_0$, then the statistic is called AOBS.

In our simulation setup, i.e. testing the null hypothesis (8) against the alternative (9), we have the point of alternative $\theta_A = (1, \theta)$ and $\theta^\star = (\theta, \theta) \in \Theta_0$ with proportions p and $1-p$, respectively. Thus $K(\theta, \theta^\star) = pK(1, \theta) + (1-p)K(\theta, \theta) = p(-\ln\theta + \theta - 1), \theta > 1$. Notice that the function $K(\theta, \theta^\star) = p(-\ln\theta + \theta - 1)$ is increasing for $\theta > 1$ and thus also the exact slope c_T is increasing with θ, since ELRH is AOBS (see Stehlík, 2003; Rublík, 1989a,b). Notice also, that the exact slope is increasing with p. It can be worth a further consideration whether this is the reason why in the lower contamination ELRH and ELR2 outperforms the other tests, at least for a large θ (see also Figure 5) and the Figure 2.

We can summarize our conclusions:

1) The power is increasing with θ.

2) The power of ELRH and ELR2 tests is relatively better for *lower contamination* than for *upper contamination*

3.4 Comparison of ELRH and ELR2 to the other tests through slopes

To obtain the slopes of other tests, we can use the Theorem of Bahadur (see Bahadur, 1967) and Groeneboom and Oosterhoff (see Groeneboom and Oosterhoff, 1977) which says that if $\lim_{n\to\infty} \frac{1}{n} \log L_n \overset{P_\theta}{\to} -\frac{1}{2}c_T(\theta) \quad \forall\, \theta \in \Theta_1$ then $N_T(\alpha, \beta, \theta) \sim \frac{2\log\frac{1}{\alpha}}{c_T(\theta)}$ for $\alpha \downarrow 0^+$. We can compute $N_T(\alpha, \beta, \theta)$ through simulation and then compare the simulated slopes of other tests to exact slopes of ELRH. Here $N_T(\alpha, \beta, \theta)$ denotes the sample size necessary for the sequence $\{T_N\}$ in order to attain power β at the level α for a point $\theta \in \Theta \setminus \Theta_0$ of the alternative space.

To get an idea of the slopes of the different tests we tried to approximate $N_T(\alpha, \beta, \theta)$. We simulated the power of each tests at level $\alpha = 0.01$ for different sample sizes $N = 10, 11, .., 20, 25, ..., 100, 200, .., 500, 1000$ and different values of θ. As an approximation to $N_T(\alpha = 0.01, \beta = 0.5, \theta)$ we used the minimum sample size $N^*(\theta)$ where the simulated power $\hat{\beta}$ was greater than 0.5. Figure 5 shows the values $c(\theta) = -2\ln\alpha/N^*(\theta)$ as a function of θ.

Note that despite the multimodal alternatives, for which a good description of test behavior can be obtained through the full variation metrics between distribution measures (see Hazan et al., 2003), in our setup the alternatives (scale exponential mixtures) are unimodal. Here we have found the Kullback-Leibler distance to be more adequate, since the ELRH test is under reasonable regularity conditions (i,ii, and iii) AOBS. Also note, that the performed simulations are not superfluous to the theoretical findings, since the justification of behavior 1) and 2) is based on

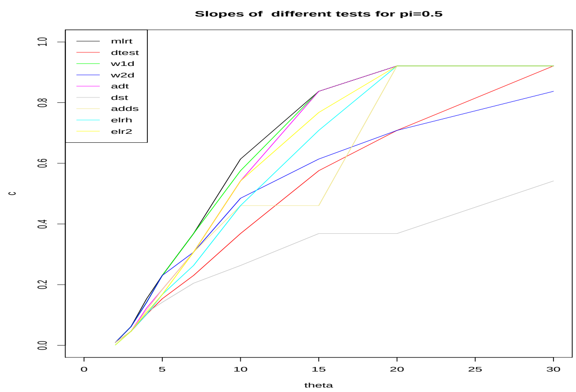

Figure 5: Slopes of different tests

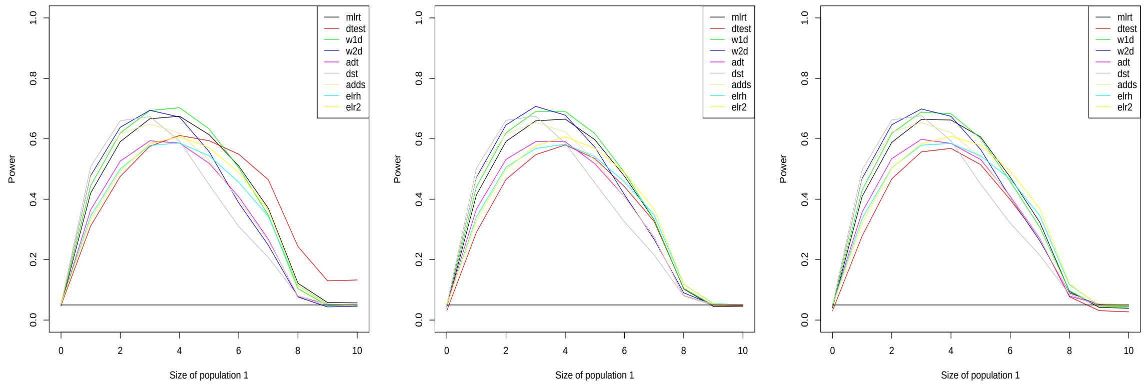

Figure 6: Proportion of rejection of the null hypothesis for different homogeneity tests for $N = 10$ and $\theta_1 = 0.1; \theta_2 = 1$ (left),$\theta_1 = 1; \theta_2 = 10$ (middle) and $\theta_1 = 10; \theta_2 = 100$ (right)

the asymptotical considerations (the nice behavior of ELRH and ELR2 tests work remarkably well also for small samples, e.g. $N = 10$ and $N = 100$ as it can be seen from the Figures).

3.5 Comparison of the tests for two subpopulations

To gain further insight into the behaviour of the different tests we conducted a simulation study for the subpopulation model. Samples of size $N = 10$ and $N = 100$ were generated from two subpopulations of size N_1 and $N - N_1$ respectively, where all integers from 0 to N were considered for N_1. We used 3 different combinations for the parameters in the two subpopulations for each sample size, namely $\theta_1 = 0.1, 1, 10$ and $\theta_2 = c\,\theta_1$. We chose $c = 10$ for $N = 10$ and $c = 3$ for $N = 100$. For each value of N_1, $M = 10000$ samples were generated and the number of rejections of the null hypothesis was counted. Figures 6 and 7 show the proportion of rejections, i.e. the simulated power of all tests as a function of the size of the first subpopulation N_1.

Note that for $N_1 = 0$ and $N_1 = N$ the sample is drawn from a homogeneous population and hence the proportion of rejections of the null hypothesis should approximately equal the size of the test $\alpha = 0.05$, which is indicated in the Figure

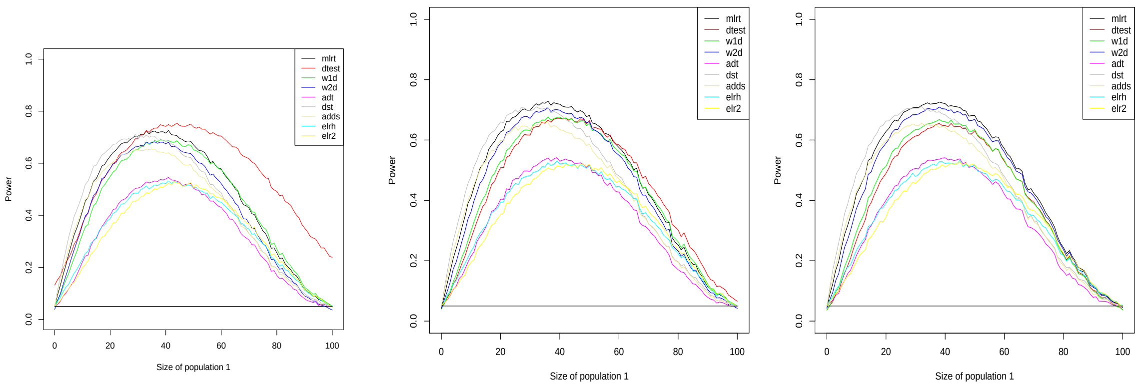

Figure 7: Proportion of rejection of the null hypothesis for different homogeneity tests for $N = 100$ and $\theta_1 = 0.1; \theta_2 = 0.3$ (left), $\theta_1 = 1; \theta_2 = 3$ (middle) and $\theta_1 = 10; \theta_2 = 30$ (right)

by the black line.

4 Illustrative Example

We have processed the inter arrival times for CAT bonds discussed thoroughly in Čížek et al. (2005). Under homogeneity the inter arrival times are iid exponential. The null hypothesis of a homogeneous exponential distribution is rejected by the ADT- and ADDS-test at significance levels $\alpha = 0.01$ and $\alpha = 0.05$ but not by all other tests (MLRT, D-test, W1D, W2D, DST, ELRH, ELR2). Thus, we can conclude in coherence with Čížek et al. (2005) that the arrival process is not a homogeneous Poisson process. It could be a renewal process or a non homogeneous Poisson process as in the CAT chapter in Čížek et al. (2005).

5 Conclusions

In the present paper we construct the efficient testing procedure of the hypotheses of homogeneity in the exponential mixture. We illustrate and explain that the ELR-tests are best for lower contamination but not for upper contamination. We also discuss the properties of such tests and describe a procedure for the computation of the critical values. We compare the performance of the exact likelihood ratio tests in the 2 component mixture alternative. In this case the ELRH can be used like the omnibus test for homogeneity, and can be, as shown in this paper, in some settings superior to other tests proposed for homogeneity in a mixture model, among them modified likelihood ratio tests, dispersion score (DS) tests ($C(\alpha)$-tests). While these approaches work well, e.g., in normal mixtures, the diagnosis of exponential mixtures poses additional problems: the modified likelihood ratio and the dispersion score tests have no power on a large class of alternatives (see Mosler and Seidel, 2001). Another widely used approach is to use a LRT statistic $2 \ln \theta = 2(l(\hat{\theta}_1) - l(\hat{\theta}_0))$ where $\hat{\theta}_0$ and $\hat{\theta}_1$ are the ML estimates of the parameters under the null and the alternative hypothesis respectively and θ denotes likelihood ratio. For this plug-in

LRT parameter estimation is usually accomplished by the EM algorithm, however the calculation of the test statistic and the Monte-Carlo simulation of its null distribution depend heavily on the particular implementation of the EM algorithm (see Seidel et al., 2000). However, this is not the case of ELRT.

The main reasons, why to use the ELRH and ELRk tests are following

a) these tests are not dependent on the unknown common scale parameter under homogeneity (like other usual tests or EM based procedures)

b) the quantiles can be easily simulated

c) these tests can be used for any sample sizes

d) the difference between the number of components, respectively, under H_0 and H_1 can be arbitrary, what is not the case of LR tests, where one looks for an asymptotical distributions

Probably the main advantages of the ELRT are simplicity of computation of test statistic and critical values for ELRH and ELR2 (e.g. no EM algorithm and simple test statistics).

Acknowledgments

The authors are grateful to Karl Mosler and Lars Haferkamp for providing the code of their simulation study and to František Rublík for his careful reading of the manuscript and helpful comments. We also thank Richard Charnigo for helpful comments concerning the D-test and its variants. Research was supported by ASO project No. SK-0607-BA-018 and by project Aktion No 50p14.

Appendix

Asymptotical optimality in the Bahadur sense

In this section we briefly discuss the asymptotical optimality in the Bahadur sense. Consider a testing problem $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_1 \setminus \Theta_0$, where $\Theta_0 \subset \Theta_1 \subset \Theta$. Further consider sequence $T = \{T_N\}$ of test statistics based on $y_1, ..., y_N$ which are iid according to an unknown member of a family $\{P_\theta : \theta \in \Theta\}$. We assume that large values of test statistics give evidence against H_0, which is the case of our test statistics of homogeneity $-\ln \Lambda_N$. For θ and t denote $F_N(t, \theta) := P_\theta\{T_N < t\}$ and $G_N(t) := \inf\{F_N(t, \theta) : \theta \in \Theta_0\}$. The random quantity $L_N = 1 - G_N(T_N)$ is called the attained level or the p-value. Suppose that for every $\theta \in \Theta_1$ the equality

$$\lim_{N\to\infty} \frac{-2 \ln L_N}{N} = c_T(\theta)$$

holds a.e. P_θ. Then the nonrandom function c_T defined on Θ_1 is called the Bahadur exact slope of the sequence $T = \{T_N\}$. According to the theorem of Raghavachari and Bahadur (see Raghavachari, 1970) the inequality

$$c_T(\theta) \leq 2K(\theta, \Theta_0) \tag{10}$$

holds for each $\theta \in \Theta_1$. Here $K(\theta, \Theta_0) := \inf\{K(\theta, \theta_0) : \theta_0 \in \Theta_0\}$ and $K(\theta, \theta_0)$ denotes the Kullback-Leibler information number. If (10) holds with the equality sign for all $\theta \in \Theta_1$, then the sequence T is said to be asymptotically optimal in the Bahadur sense. The maximization of $c_T(\theta)$ is a nice statistical property. The class of such statistics is apparently narrow, though it contains under certain conditions the LR statistics (see Bahadur, 1965, 1967; Rublík, 1989a,b). Rublík proved AO of the LR statistic under regularity condition which is shown to be fullfiled by regular normal, exponential and Laplace distribution under additional assumption that Θ_0 is a closed set and Θ_1 is either closed or open in metric space Θ. For more extensive discussion on asymptotical optimality see also the monograph Nikitin (1995).

References

Bahadur, R. (1965). An optimal property of the likelihood ratio statistic. In *Proc.5th Berkeley Sympos. on Probab. Theory and Mathem. Statist., vol. 1*, Berkeley and Los Angeles, pp. 13–26. Univ. of California Press.

Bahadur, R. (1967). Rates of convergence of estimates and test statistics. *The Annals of Mathematical Statistics 38*, 303–324.

Balakrishnan, N. and A. P. Basu (1996). *The Exponential Distribution: Theory, Methods, and Applications.* New York: Gordon and Breach.

Charnigo, R. and J. Sun (2004). Testing homogeneity in a mixture distribution via the l^2 distance between competing models. *Journal of the American Statistical Association 99*, 488 – 498.

Chen, H., J. Chen, and J. D. Kalbfleisch (2001). A modified likelihood ratio test for homogeneity in finite mixture models. *Journal of Royal Statistical Society, Series B 63*, 19 – 29.

Choi, S. (1979). Two-sample tests for compound distributions for homogeneity of mixing proportions. *Technometrics 21*(3), 361–365.

Čížek, P., W. Härdle, and R. Weron (2005). *Statistical Tools in Finance and Insurance. ISBN 3-540-22189-1.* Heidelberg: Springer Verlag.

Efimova, T., V. Leskin, G. Ososkov, K. Tolstov, and N. Chernov (1989). Expansion of transverse momenta in inelastic collisions of particles into rayleigh distributions. *JINR Rapid Communications 3[36].*

Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models.* New York: Springer-Verlag:Springer Series in Statistics.

Garel, B. (2007). Recent asymptotic results in testing for mixtures. *Computational Statistics and Data Analysis 51*, 5295–5304.

Groeneboom, P. and J. Oosterhoff (1977). Bahadur efficiency and probabilities of large deviations. *Statist. Neerlandica 31*, 1–24.

Hartigan, J. (1985). A failure of likelihood ratio asymptotics for normal mixtures. In *proc. Berkeley conference in honor of Jerzy Neyman and Jack Kiefer*, Monterey, CA. Wasworth Advanced books.

Hazan, A., Z. Landsman, and U. Makov (2003). Robustness via a mixture of exponential power distributions. *Computational Statistics and Data Analysis 42*, 111–121.

Henze, N. and S. Meintanis (2005). Recent and classical tests for exponentiality: a partial review with comparisons. *Metrika 61*, 29–45.

Lehmann, E. (1964). *Testing Statistical Hypotheses*. New York: John Wiley & Sons.

Lindsay, B. (1995). *Mixture models: Theory, geometry and applications*. Hayward, Cal.: Institute of Mathematical Statistics.

Liu, X., C. Pasarica, and Y. Shao (2003). Testing homogeneity in gamma mixture models. *Scand. J. Statist. 30*, 227–239.

Manoukian, E. (1986). *Modern concepts and theorems of mathematical statistics*. New York: Springer-Verlag:Springer Series in Statistics.

McLachlan, G. (1995). Mixtures - models and applications. the exponential distribution. In *proc. Berkeley conference in honor of Jerzy Neyman and Jack Kiefer*, Amsterdam, pp. 307–323. Gordon & Breach.

Meintanis, S. (2007). Test for exponentiality against weibull and gamma decreasing hazard rate alternatives. *KYBERNETIKA 43*(3), 307–314.

Mosler, K. and L. Haferkamp (2007). Size and power of recent tests for homogeneity in exponential mixtures. *Communications in Statistics - Simulation and Computation 36*, 493–504.

Mosler, K. and W. Seidel (2001). Testing for homogeneity in an exponential mixture model. *Australian and New Zealand Journal of Statistics 43*, 231–247.

Nikitin, Y. (1995). *Asymptotic Efficiency of Nonparametric Tests*. New York: Cambridge University Press.

Raghavachari, M. (1970). On a theorem of bahadur on the rate of convergence of test statistics. *Ann. Mathem. Statist. 41*, 1695–1699.

Randles, R. H. (1982). On the asymptotic normality of statistics with estimated parameters. *Ann. Statist. 10*, 462–474.

Rublík, F. (1989a). On optimality of the LR tests in the sense of exact slopes, part 2, application to individual distributions. *Kybernetika 25*, 117–135.

Rublík, F. (1989b). On optimality of the LR tests in the sense of exact slopes,Part 1, general case. *Kybernetika 25*, 13–25.

Seidel, W., K. Mosler, and M. Alker (2000). A cautionary note on likelihood ratio tests in mixture models. *Annals of the Institute of Statistical Mathematics 52*, 481–487.

Stehlík, M. (2003). Distributions of exact tests in the exponential family. *Metrika 57*, 145–164.

Stehlík, M. (2006). Exact likelihood ratio scale and homogeneity testing of some loss processes. *Statistics and Probability Letters 76*, 19–26.

Stehlík, M. (2008). Homogeneity and scale testing of generalized gamma distribution. *Reliability Engineering & System Safety 93*, 1809–1813.

Stehlík, M. and G. Ososkov (2003). Efficient testing of the homogeneity, scale parameters and number of components in the rayleigh mixture. *JINR Rapid Communications E-11-2003-116*.

Susko, E. (2003). Weighted tests of homogeneity for testing the number of components in a mixture. *Computational Statistics and Data Analysis 41*, 367–378.

Tchirina, A. (2005). Large deviations for a class of scale-free statistics under the gamma distribution. *Journal of Mathematical Sciences 128(1)*, 2640–2655.