



HAL
open science

Exact likelihood ratio testing for homogeneity of exponential distribution

Milan Stehlik, Helga Wagner

► **To cite this version:**

Milan Stehlik, Helga Wagner. Exact likelihood ratio testing for homogeneity of exponential distribution. *Communications in Statistics - Simulation and Computation*, 2011, 40 (05), pp.663-684. 10.1080/03610918.2010.551011 . hal-00680009

HAL Id: hal-00680009

<https://hal.science/hal-00680009>

Submitted on 17 Mar 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Exact likelihood ratio testing for homogeneity of exponential distribution

Journal:	<i>Communications in Statistics - Simulation and Computation</i>
Manuscript ID:	LSSP-2009-0168.R2
Manuscript Type:	Original Paper
Date Submitted by the Author:	07-Jun-2010
Complete List of Authors:	Stehlik, Milan; Johannes Kepler University, Department of Applied Statistics Wagner, Helga; Johannes Kepler University
Keywords:	Homogeneity testing, exponential distribution, likelihood ratio, subpopulation model, mixture models, asymptotic efficiency
Abstract:	The aim of this paper is to discuss homogeneity testing of the exponential distribution. We introduce the exact likelihood ratio test of homogeneity in the subpopulation model, ELR, and the exact likelihood ratio test of homogeneity against the two-components subpopulation alternative, ELR2. The ELR test is asymptotically optimal in the Bahadur sense when the alternative consists of sampling from a fixed number of components. Thus in some setups the ELR is superior to frequently used tests for exponential homogeneity which are based on the EM algorithm. We demonstrate this fact by both theoretical comparisons and simulations.
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
StehlikWagner.rar	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

SCHOLARONE™
Manuscripts

For Peer Review Only

Exact Likelihood ratio testing for homogeneity of the exponential distribution

Milan Stehlík and Helga Wagner

Department of Applied Statistics, Johannes Kepler University Linz,
Freistädter Straße 315, A-4040 Linz, Austria
milan.stehlik@jku.at, helga.wagner@jku.at

May 26, 2010

Key Words: Homogeneity testing; exponential distribution; likelihood ratio; subpopulation model; exact distribution; asymptotic efficiency; exact slopes; mixture models.

Abstract

The aim of this paper is to **discuss** homogeneity testing of the exponential distribution. We introduce the exact likelihood ratio test of homogeneity in the subpopulation model, ELR, and the exact likelihood ratio test of homogeneity **against the two-components subpopulation alternative**, ELR2. The ELR test is asymptotically optimal in the Bahadur sense when the alternative consists of sampling from a fixed number of components. **Thus in some setups the ELR is superior to frequently used tests for exponential homogeneity which are based on the EM algorithm (like the MLRT, the ADDS test, and the D-tests).** One important example of superiority of ELR and ELR2 tests is the case of lower contamination. We demonstrate this fact by both theoretical comparisons and simulations.

1 Introduction

The exponential distribution $Exp(\theta_i)$ 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 and a constant hazard rate θ_i . 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).

Lifetime data with an observed decreasing hazard rate rate can be modelled by a mixture distribution of exponential components. Such a model can be interpreted 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 exact procedure for testing exponential homogeneity against alternatives of exponential heterogeneity. The likelihood-ratio decision procedure for 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 and requires specification of the null and the alternative.

When we consider the null hypothesis of exponential homogeneity, different specifications of the alternative lead to different likelihood ratio tests. The most popular alternative to homogeneity is the mixture model with exponential components. A certain drawback of the LR test for the mixture model alternative is that it requires ML estimates of the model parameters. This is not the case when the alternative is the subpopulation model, which is the alternative of the exact likelihood ratio (ELR) tests considered in this paper.

We compare these procedures with other competitive tests and show, in particular, its superiority in some setups (e.g. in the lower contamination case, which can be of interest for insurance applications, among others).

1 The ELR-test is asymptotically optimal in the Bahadur sense when the alternative consists
2 of the subpopulation model with a finite number of populations (see Stehlík, 2006; Rublík,
3 1989a,b). The ELR test and ELR2 test, which tests the alternative of two subpopulations, have
4 nonstandard asymptotic distributions but their exact distribution can be easily simulated.
5
6
7

8 The rest of this paper is organized as follows. The difference between mixture and subpopu-
9 lation models is discussed in Section 2. In section 3 the exact likelihood ratio homogeneity tests
10 ELR and ELR2 are introduced and discussed. In section 4 a comparative power study of tests
11 for homogeneity is provided together with the theoretical explanation of the obtained results.
12 The following tests for homogeneity in exponential mixtures are considered: the dispersion
13 score (DS) test, also known under the name $C(\alpha)$ -tests; (see chapter 4 of Lindsay (1995); the
14 modified likelihood ratio test (MLRT) introduced by Chen et al. (2001), which is a penalized
15 LRT and has standard chi-square asymptotics; the ADDS test by Mosler and Seidel (2001), a
16 combination ... procedure; several variants of the D-test by Charnigo and Sun (2004), that is
17 based on the L^2 distance between the estimated densities of a homogeneous and a heterogeneous
18 model. All tables and proofs are collected in the Appendix.
19
20
21
22
23
24
25
26
27
28
29
30

31 2 The subpopulation model and mixture alternative

32 2.1 Subpopulation and mixture models

33 Consider a sample of independent observations y_1, \dots, y_N , and the null hypothesis of exponen-
34 tial homogeneity, that is
35
36

$$37 f(y_1, \dots, y_N) = \prod_{i=1}^N \theta \exp(-\theta y_i)$$

38 For testing this null hypothesis, many different alternatives may be specified. The most general
39 alternative says that the distributions of the y_i are heterogeneous and/or non-exponential, which
40 comprises alternatives like homogeneous and non-exponential, heterogeneous and exponential
41 in subgroups, heterogeneous and non-exponential in at least one subgroup, among others. In
42 the sequel, we investigate the following alternative: each observation y_i is generated by an
43 exponential distribution but the parameter is not identical for all observations.
44
45
46
47
48
49
50
51
52
53

54 Alternatives to homogeneity are often specified as mixture models. The joint density of a
55
56
57
58
59
60

sample y_1, \dots, y_N of iid observations from a 2-component mixture is given as

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

and the joint density of a sample from a general k - component mixture of exponential components is

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

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

In this paper we consider tests for homogeneity against somewhat different alternatives, which are specified as subpopulation models. In the subpopulation model the number of subpopulations has to be specified. The general subpopulation model which is the alternative tested in the exact likelihood ratio test for homogeneity (ELR) proposed by Stehlík (2003) assumes that each observation follows an exponential distribution with its own parameter. The joint density of the sample is given as

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

$\theta_i \neq \theta_j$ for $i \neq j$.

A more specific case of a subpopulation model is inhomogeneity with an unobserved clustering and a given number of clusters k , this is the alternative of the exact likelihood ratio test for k subpopulations ELR $_k$ introduced in (Stehlík and Ososkov, 2003). The ELR $_2$ test uses the alternative of two subpopulations, which can be specified by (1) and the existence of two nonempty index sets M_1, M_2 such that

$$M_1 \cup M_2 = \{1, \dots, N\}, M_1 \cap M_2 = \emptyset \quad (2)$$

$$\forall j \in M_1 : \theta^j = \theta^1, \forall j \in M_2 : \theta^j = \theta^2, \theta^1 \neq \theta^2. \quad (3)$$

An even more restrictive alternative within the subpopulation model is inhomogeneity with a fixed number of clusters and observed clustering (see Hönig et al. (2010)). Such a situation may appear in some specific, e.g. biological applications.

1 Both the mixture and the subpopulation model can be used to model unobserved cluster-
2 ing. To illustrate the difference between these models we consider the related data generating
3 processes for sampling from two exponential distributions $Exp(\theta_i), i = 1, 2; \theta_1 \neq \theta_2$. A scheme
4 where the number of draws $N_1, 0 < N_1 < N$, from $Exp(\theta_1)$ is fixed before sampling can be
5 described by the subpopulation model. If the component $Exp(\theta_i)$ from which each element of
6 the sample is drawn is determined by a random experiment the mixture model results. Under
7 H_1 in the 2-subpopulation model the sample contains at least one element of each compo-
8 nent, whereas this need not be the case for samples from the two-component mixture model.
9 We might therefore interpret the alternative in the subpopulation model as exponential het-
10 erogeneity in the sample, whereas the mixture alternative can be interpreted as exponential
11 heterogeneity in the population. The mixture alternative is more commonly used in testing
12 homogeneity, but the subpopulation alternative appears to be a useful surrogate.

13
14
15
16
17
18
19
20
21
22
23 The reason why we consider the subpopulation model is, besides simplicity, the fact that as
24 soon the difference between the number of components in the mixture model under H_0 and H_1
25 respectively is greater than 1, the likelihood ratio tests involves nonstationary random fields, for
26 which very few theoretical results are available (see Garel, 2007). The ELR-test is asymptoti-
27 cally optimal in the Bahadur sense when the alternative consists of the subpopulation model
28 with a finite number of populations (see Stehlik, 2006; Rublik, 1989a,b). ELR and ELR2 tests
29 have nonstandard asymptotic distributions but their exact distribution can easily be simulated.
30 Generalization to testing for k subpopulations by the ELR k test is easy to implement, however
31 simulation of critical values is computationally expensive and computational cost increases with
32 k , as a min or max over all possible partitions of a sample (y_1, \dots, y_n) into k subsets has to
33 be computed. Application of permutation principles will be worth further investigation. Some
34 applications of ELR k can be found in physics, see e.g. Efimova et al. (1989).

35
36
37
38
39
40
41
42
43
44
45 Our setup encompasses also the case of the Weibull distribution with known shape pa-
46 rameter. Tests for homogeneity when the shape parameter of the Weibull is unknown have
47 been developed by Mosler and Scheicher (2008); for a comparison of procedures, see Mosler
48 and Haferkamp (2009). Tests for exponentiality against a Weibull alternative are given in
49 Meintanis (2007) and Henze and Meintanis (2005).

2.2 Asymptotic approaches to testing for homogeneity

When the sample size is increasing, we consider the following setup for the m -subpopulation model. The set of overall parameters Θ consists of the m -tuples $\theta = (\theta_1, \dots, \theta_m)$, where θ_j is the parameter of the j -th population. Let us suppose that in the k -th experiment the size of the sample from the j -th population is $n_k^{(j)}$, $j = 1, \dots, m$ and $k = 1, 2, \dots$. Let the product measure $P_{\theta_j}^\infty$ correspond 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 $P_\theta = P_{\theta_1}^\infty \times \dots \times P_{\theta_m}^\infty$, can be used to describe the 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 the ELR test:

- i) if $k \neq l$ then $n_k^{(j)} \neq n_l^{(j)}$ for some j
- ii) $\lim_{k \rightarrow \infty} n_k = +\infty$
- iii) $\lim_{k \rightarrow \infty} \frac{n_k^{(j)}}{n_k} = p_j \in (0, 1]$, $j = 1, \dots, m$

For $\theta = (\theta_1, \dots, \theta_m)$, $\theta^* = (\theta_1^*, \dots, \theta_m^*) \in \Theta$ let $K(\theta, \theta^*) = \sum_{j=1}^m p_j K(\theta_j, \theta_j^*)$, where the Kullback-Leibler information is defined by the formula

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

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

For the alternative of the two-component mixture form, i.e. for the testing problem

$$H_0 : \theta_0 \exp(-\theta_0 x) \text{ versus } H_A : p\theta_0 \exp(-\theta_0 x) + (1-p)\theta \exp(-\theta x), \theta > \theta_0, 0 < p \leq 1, \quad (4)$$

where θ_0 is known and θ, p are unknown parameters we will get a similar asymptotical behavior as Hartigan (1985). He discovered the divergence of the LR test statistics for homogeneity testing in normal mean mixture models with an unbounded mean parameter. The classical chi-squared limiting distributional result of Wilks (1938) is not applicable. For exponential mixtures, the likelihood ratio test converges in distribution to the sup of a square

Gaussian process (see Ciuperca (2002)). Liu et al. (2003) have proved in the setup (4) that $\lim_{N \rightarrow \infty} P(2\Lambda_N - \ln \ln N + \ln(16\pi^2) \leq x) = \exp(-\exp(-x/2))$. They also try to determine whether it is feasible to approximate $2\Lambda_N - \ln \ln N + \ln(16\pi^2)$ 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. The difficulties with regularity conditions under the homogeneity hypothesis have been recently investigated by Li et al. (2009).

3 Exact Likelihood Ratio Tests of Homogeneity

3.1 The ELR test

For a sample of N independent observations $y = (y_1, \dots, y_N)$, where $y_i \sim \text{Exponential}(\theta_i)$ we consider the LR homogeneity test against the alternative (1) (subpopulation model), i.e.

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

The exact distribution of the LR test for homogeneity against the alternative (1), the ELR 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 are optimal in regular cases.

The LR statistics has the form (see Theorem 3 of Stehlík (2006))

$$-\ln \Lambda_N = N \ln\left(\sum_{i=1}^N y_i\right) - N \ln N - \sum_{i=1}^N \ln y_i. \quad (5)$$

A very important property of the LR test for homogeneity is its scale invariance, i.e. the distribution of the test statistic under H_0 is independent of the unknown scale parameter. This is an advantage compared to some asymptotic tests and tests depending on the true but unknown value of θ . The critical values $c_{1-\alpha}$ are easy to obtain by simulation, e.g. from the standard exponential 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, \dots, c_M for the test statistic (5) under homogeneity. The values $c_{1-\alpha}$ are determined as the respective order statistic $c_{1-\alpha} = c_{(M(1-\alpha))}$. These values are used throughout the paper.

Table 1: Critical values for the ELR

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

Under homogeneity the statistics (5) is a monotonous function of the statistic $\left(\frac{N}{y_1 + \dots + y_N}\right)^N \prod_{i=1}^N y_i$ which is clearly scale independent. The LR homogeneity statistic (5) is also a monotonous transformation of the so called Moran's statistics T_N^+ and T_N^- , where

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

$T_N^- = -T_N^+$, $C \approx 0.57721566$ is the Euler constant and \bar{y} is the arithmetic mean. This is a scale-free **exponentiality test that has been studied in** Tchirina (2005). The ELR test for homogeneity provided in this paper is 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).

3.1.1 Simulation Study

A simulation study was **carried out** 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 $M = 100000$ samples were generated and the proportion of rejections of the ELR test was determined.

Results given in Table 2 show that the ELR test holds the chosen size α also for small samples. The power of the ELR 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 ELR 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 $p = 0.9$, 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. For a given θ lower contamination ($p = 0.9$) is harder to detect than upper contamination ($p = 0.1$) as overdispersion measured by the squared coefficient of variation is higher for upper contamination

3.2 The ELR2 test

In this section we will discuss the efficient testing procedure of the number of components k in the exponential mixture for $k = 2$, introduced by Stehlík and Ososkov (2003). Here we consider LR homogeneity testing with a more complex alternative H_1 , which is the approximation to a finite scale mixture. In physics, such testing corresponds to the testing of the number k of secondary particles obtained after the collision: Homogeneity corresponds to one particle ($k = 1$), and $k > 1$ corresponds to k particles (cf Efimova et al., 1989; Stehlík and Ososkov, 2003). We consider the alternative of the form $H_1 : k = 2$. The hypothesis

$$H_0 : k = 1 \text{ versus } H_1 : k = 2 \quad (6)$$

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

$$H_0 : \theta_1 = \dots = \theta_n \text{ versus } H_1 : \exists \text{ non-empty } M_1, M_2, M_1 \cup M_2 = \{1, \dots, N\}, \quad (7)$$

where (2) and (3) hold. We construct the LR test of the hypothesis (7) which approximates the hypothesis (6).

Table 2: Simulated power for the ELR-test

θ	1	2	3	4	5	6	7	8	9	10
$\alpha = 0.05$										
$p = 0.1$										
N=20	0.0548	0.0665	0.1200	0.1532	0.2459	0.2874	0.3337	0.3797	0.4328	0.4704
N=50	0.0455	0.0734	0.1594	0.2844	0.4038	0.5070	0.6012	0.6479	0.7265	0.7470
N=100	0.0596	0.0980	0.2331	0.4174	0.5860	0.6976	0.8123	0.8660	0.9081	0.9300
N=1000	0.0562	0.2871	0.8401	0.9941	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$p = 0.5$										
N=20	0.0470	0.0906	0.1844	0.3077	0.3793	0.5057	0.5842	0.6543	0.7383	0.7618
N=50	0.0547	0.1243	0.3125	0.5339	0.7271	0.8338	0.9196	0.9546	0.9750	0.9921
N=100	0.0514	0.1738	0.5017	0.7740	0.9357	0.9797	0.9982	0.9970	1.0000	0.9994
N=1000	0.0570	0.7254	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$p = 0.9$										
N=20	0.0561	0.0617	0.0723	0.1003	0.1109	0.1200	0.1378	0.1579	0.1713	0.1800
N=50	0.0448	0.0787	0.0950	0.1293	0.1575	0.1886	0.2211	0.2496	0.2805	0.2942
N=100	0.0490	0.0807	0.1176	0.1851	0.2288	0.3100	0.3498	0.3950	0.4362	0.4632
N=1000	0.0517	0.1826	0.4694	0.7505	0.8980	0.9667	0.9881	0.9961	0.9993	0.9998
$\alpha = 0.01$										
$p = 0.1$										
N=20	0.0100	0.0148	0.0333	0.0600	0.0896	0.1502	0.1876	0.2261	0.2577	0.3086
N=50	0.0100	0.0173	0.0437	0.1223	0.2205	0.3167	0.4105	0.5180	0.5619	0.6407
N=100	0.0114	0.0314	0.0844	0.2162	0.3838	0.5412	0.6382	0.7607	0.8048	0.8699
N=1000	0.0114	0.1063	0.6518	0.9718	0.9994	1.0000	1.0000	1.0000	1.0000	1.0000
$p = 0.5$										
N=20	0.0098	0.0225	0.0642	0.1053	0.1718	0.2523	0.3255	0.3891	0.4555	0.5232
N=50	0.0126	0.0286	0.1172	0.2680	0.4778	0.6055	0.7537	0.8414	0.8891	0.9375
N=100	0.0102	0.0651	0.2288	0.5462	0.7764	0.9062	0.9723	0.9903	0.9976	0.9982
N=1000	0.0103	0.4746	0.9992	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$p = 0.9$										
N=20	0.0141	0.0135	0.0184	0.0213	0.0262	0.0298	0.0389	0.0433	0.0461	0.0517
N=50	0.0080	0.0160	0.0213	0.0379	0.0446	0.0619	0.0693	0.0886	0.1014	0.1193
N=100	0.0103	0.0229	0.0374	0.0627	0.0844	0.1149	0.1379	0.1716	0.2074	0.2461
N=1000	0.0120	0.0575	0.2211	0.4998	0.7301	0.8692	0.9450	0.9766	0.9893	0.9957

Let y_1, \dots, y_N be independently distributed with exponential densities. Suppose that $\{y_{i_1}, \dots, y_{i_K}\}$, $0 < K < N$ are the observations from the exponential distribution with scale parameter θ_1 and the other observations are distributed according to the exponential distribution with scale parameter θ_2 . Then following Stehlík and Ososkov (2003) we obtain the formula for likelihood ratio

$$\Lambda_N(y) = \min_{0 < K < N, p \in P(K)} \left\{ \frac{N^N}{K^K (N - K)^{N-K}} \frac{(y_{i_1} + \dots + y_{i_K})^K (y_{i_{K+1}} + \dots + y_{i_N})^{N-K}}{(y_1 + \dots + y_N)^N} \right\}, \quad (8)$$

where $P(K)$ denotes all partitions of $\{1, \dots, K\}$ in two nonempty subsets. The main advantage of the test statistic (8) is that it does not depend on the unknown value of the parameter θ under the null hypothesis. The **exact** distribution of the LR test statistic Λ_N given by formula (8) is derived in Stehlík and Ososkov (2003) under the null hypothesis.

The following Lemma provides some notes on the determination of the ELR2 test statistics. The proof is given in the Appendix 6.1.

Lemma 1 *The likelihood-ratio statistic of the ELR2 test*

$$-\ln \Lambda_N(y) = -\min_{0 < K < N, p \in P(K)} \{ N \ln N - K \ln K - (N - K) \ln(N - K) + K \ln(\sum_{n=1}^K y_{i_n}) + (N - K) \ln(\sum_{n=1}^{N-K} y_{i_n}) - N \ln(\sum_{n=1}^N y_n) \},$$

can be determined as

$$\ln \Lambda_N(y) = N \ln N - N \ln \left(\sum_{n=1}^N y_n \right) + H_{\min},$$

where

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

H_{\min} can be determined very simply as only sums of order statistics are involved. Critical values can be obtained similarly to the ELR 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 $c_{1-\alpha} = c_{(M(1-\alpha))}$.

4 A comparative power study of **different** homogeneity tests

4.1 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 ELR and ELR2 to other different tests for the hypotheses

$$H_0 : y_1, \dots, y_N \sim \text{Exponential}(\theta) \quad (9)$$

against the alternative

$$H_1 : y_1, \dots, y_N \text{ follow a mixture of two exponential components} \quad (10)$$

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 $N_1 \sim \text{Binomial}(N, p)$. We used three typical mixture proportions, $p = 0.1, 0.5, 0.9$ and the sample sizes $N = 10, 100$ and $N = 1000$. The parameter of the first mixture component is $\theta_1 = 1$ and for θ_2 several values greater than θ_1 were 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 two 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 Seidel (2001) (ADDS) and the exact likelihood ratio tests against the alternative (1) (ELR) and for testing against the 2-component subpopulation model (ELR2).

We also made comparisons to the 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 personally communicated by Richard Charnigo, the

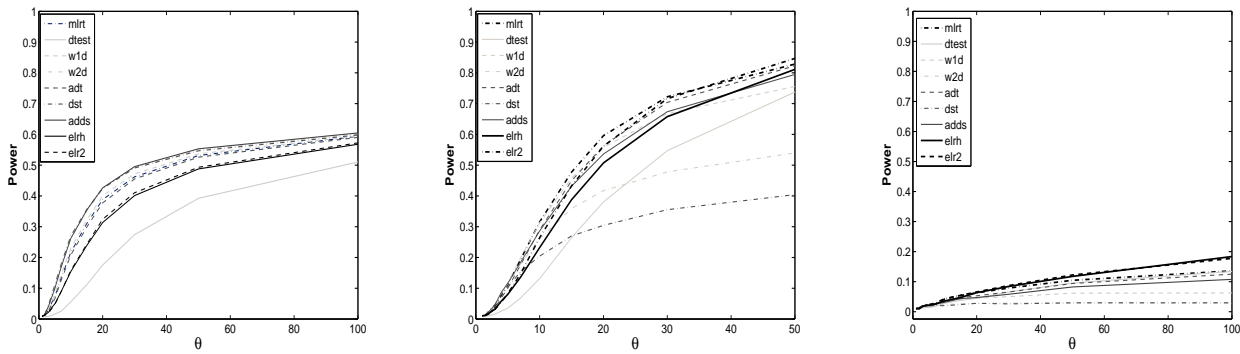


Figure 1: Proportion of rejections of the null hypothesis for different homogeneity tests for $N = 10$, $\alpha = 0.05$ and $p = 0.1$ (left), $p = 0.5$ (middle) and $p = 0.9$ (right)

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 the scale invariance of the ELR and the ELR2 test.

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 values for the AD and the DS-test.

4.2 Results

The simulated powers of the different homogeneity tests ($\alpha = 0.01$) of exponential mixtures are reported in Tables 3– 5 (see Appendix 6.2) ordered by decreasing powers. In Figure 1 the results are displayed for $N = 10$ and Figure 2 shows the results in the lower contamination case for $N = 100$ and $N = 1000$.

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 ELR-tests are among those with lowest power, with ELR performing better than ELR2. Only the D-test has smaller power than ELR 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 $N = 1000$),

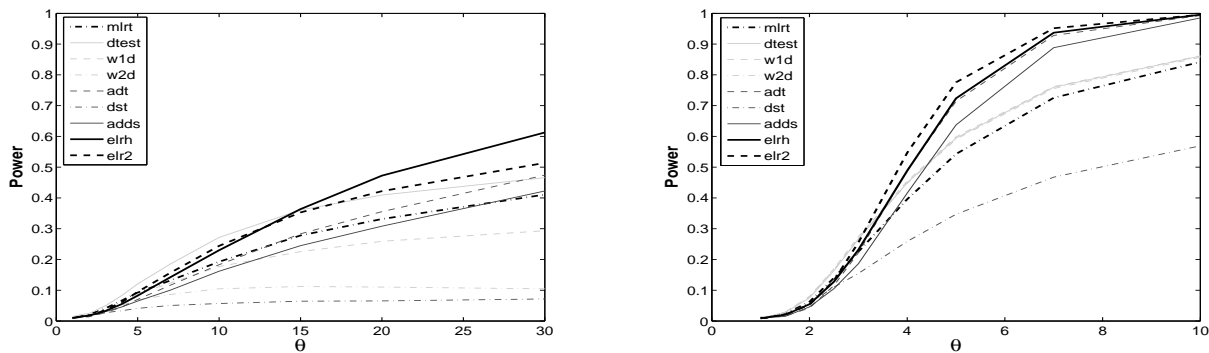


Figure 2: Proportion of rejections of the null hypothesis for different homogeneity tests for $\alpha = 0.01$, $p = 0.1$ and $N = 100$ (left), and $N = 1000$ (right)

but only for smaller values of θ for $N = 10$. In small samples of mixtures with $\theta > 20$ the ELR2-test is the one attaining the highest power, but in larger samples the ELR 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 sizes. In this situation however the ELR 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 behavior as a consequence of the 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.

4.3 Theoretical explanation of the test behavior

In our simulation setup, i.e. testing the null hypothesis (9) against the alternative (10), we have the point of alternative $\theta_A = (1, \theta)$ and $\theta^* = (\theta, \theta) \in \Theta_0$ with proportions p and $1 - p$, respectively. Thus $K(\theta, \theta^*) = pK(1, \theta) + (1 - p)K(\theta, \theta) = p(-\ln \theta + \theta - 1)$, $\theta > 1$. Notice that the function $K(\theta, \theta^*) = p(-\ln \theta + \theta - 1)$ is increasing for $\theta > 1$ and thus also the exact slope c_T is increasing with θ , since ELR is AOBS (see Stehlík, 2003; Rublík, 1989a,b). Notice also, that the exact slope is increasing with p . It could be worth further consideration whether this

is the reason why in the lower contamination ELR and ELR2 outperforms the other tests, at least for large θ (see Figure 1 and Figure 3).

We can summarize our conclusions:

- 1) The power is increasing with θ .
- 2) The power of ELR and ELR2 tests is relatively better for *lower contamination* than for *upper contamination*.

4.4 Comparison of ELR and ELR2 to the other tests through slopes

The interpretation of the behavior of the simulated power can be based on exact slopes of the sequence of tests. 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 $\lim_{n \rightarrow \infty} \frac{1}{n} \ln L_n \xrightarrow{P_\theta} -\frac{1}{2}c_T(\theta) \quad \forall \theta \in \Theta_1$ then $N_T(\alpha, \beta, \theta) \sim \frac{2 \ln \frac{1}{\alpha}}{c_T(\theta)}$ 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 ELR. 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 test at level $\alpha = 0.01$ for different sample sizes $N = 10, 11, \dots, 20, 25, \dots, 100, 200, \dots$ 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 3 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 ELR test is AOBS under reasonable regularity conditions (i,ii, and iii). Also note, that the performed simulations are not superfluous to the theoretical findings, since the justification of behavior 1) and 2) is based on the asymptotical considerations (the nice behavior of ELR 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).

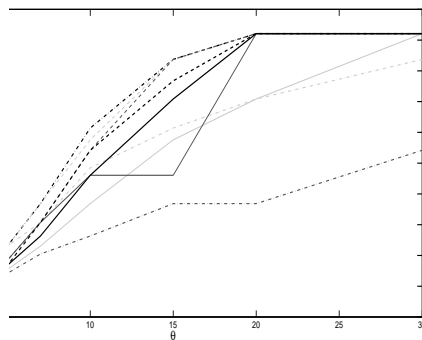


Figure 3: Slopes of different tests

4.5 Comparison of the tests for two subpopulations

To gain further insight into the behavior of the different tests for various modalities of the alternative we conducted a further simulation study. 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 . The difference to the simulation in Section 4.1 where the number of observations from the first mixture component N_1 was generated from a binomial distribution, is that N_1 is fixed.

We used three 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. Figure 4 shows the proportion of rejections, i.e. the simulated power of all tests as a function of the size of the first subpopulation N_1 for $N = 10$.

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 by a black line in Figure 4.

This simulation study provides some further insight in the properties of the different tests:

1. The D-Test obviously is not scale invariant under the null see Figure 4 right. This
2. The power of all tests is not symmetric in $|N_1 - N/2|$ and is small for lower contamination.
3. No test dominates the others.

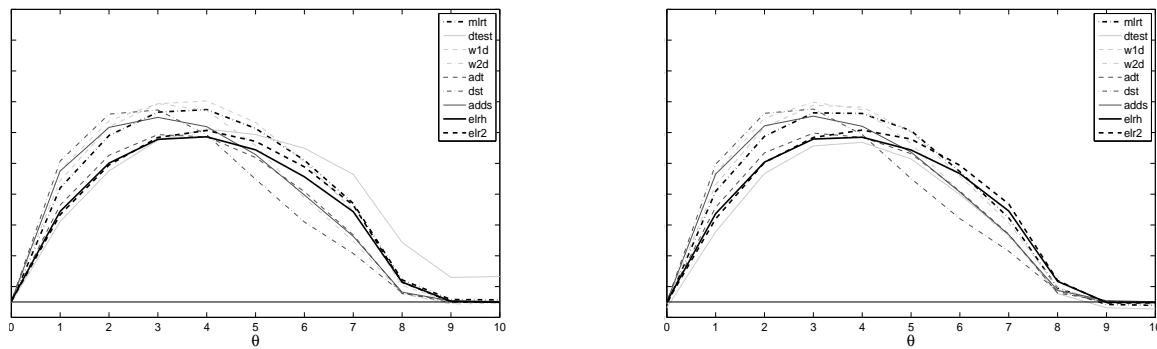


Figure 4: Proportion of rejection of the null hypothesis for different homogeneity tests for $N = 10$ and $\theta_1 = 0.1; \theta_2 = 1$ (left) and $\theta_1 = 10; \theta_2 = 100$ (right)

4. For fixed N and N_1 the power seems – at least approximately – to be a function of θ_1/θ_2 .

5 Conclusions

In the present paper we construct the efficient testing procedure for exponential homogeneity. The ELR procedures studied here have some interesting properties under subpopulation alternatives, while in our simulations we employ these tests against mixture component alternatives and compare them with other existing tests developed in the mixture component model. We illustrate by simulations and try to explain also theoretically 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 ELR can be used like the omnibus test for homogeneity and in some settings can be superior to other tests proposed for homogeneity in a mixture model, among them modified likelihood ratio tests or dispersion score (DS) 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). For the tests used in this paper, in the software implementation D-Test, w1d, w2d and MLRT tests use estimates from the EM Algorithm. The AD-test uses a scale estimate (mean) under the null. However, this is not the case of ELR and ELRk tests.

The main reasons, why the ELR and ELRk tests should be considered for testing exponential homogeneity are the 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.

Probably the main advantage of the ELRT is its computational simplicity for ELR and ELR2, regarding both the test statistic and the critical values; in particular, no EM algorithm has to be employed.

Acknowledgments

The authors are thankful for the constructive comments of two anonymous referees, which substantively improved the quality of the paper. 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 54p13.

6 Appendix

6.1 Proofs

Proof of Lemma 1 The determination of the likelihood-ratio statistic of the ELR2 test

$$-\ln \Lambda_N(y) = -\min_{0 < K < N, p \in P(K)} \{N \ln N - K \ln K - (N - K) \ln(N - K) + \\ + K \ln(\sum_{n=1}^K y_{i_n}) + (N - K) \ln(\sum_{n=1}^{N-K} y_{i_n}) - N \ln(\sum_{n=1}^N y_n)\},$$

is not so straight forward as for the ELR, as the minimum of

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

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

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

is irrelevant and therefore the minimum of

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

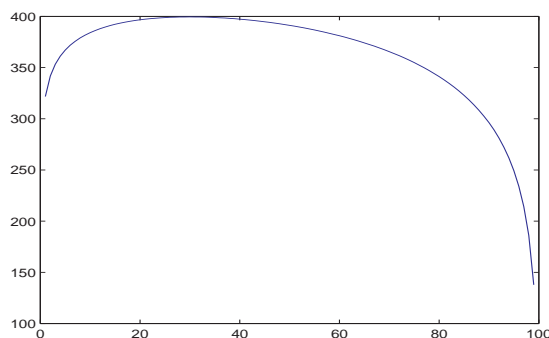
for $0 < K < N, p \in P(K)$ is of interest. For N observations there are $2^{N-1} - 1$ different classifications into 2 nonempty groups. Hence direct minimization over all classifications is feasible only for small N . However minimizing first $H(y|K)$ for fixed $K = 1, \dots, N - 1$ and then determining 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

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

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 5. For the likelihood ratio statistic, x can

Figure 5: The function $H(x|K)$

only take certain discrete values, as $x = \sum_{n=1}^K y_{i_n}$. Minimum and maximum of x are the sum of the smallest and largest K order statistics respectively:

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

Due to the symmetry

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

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

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

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

$$\ln \Lambda_N(y) = N \ln N - N \ln\left(\sum_{n=1}^N y_n\right) + H_{\min}.$$

□

6.2 Tables

References

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

Table 3: Simulated power for different tests for homogeneity in exponential mixtures $n = 10$

Rank	Test	Power	Test	Power	Test	Power	Test	Power	Test	Power
$p = 0.1$										
		$\theta = 1$		$\theta = 5$		$\theta = 10$		$\theta = 20$		$\theta = 50$
1	adds	(0.010)	dst	(0.111)	dst	(0.270)	adds	(0.426)	adds	(0.553)
2	w1d	(0.010)	adds	(0.099)	adds	(0.262)	dst	(0.425)	w2d	(0.549)
3	elr2	(0.009)	w2d	(0.096)	w2d	(0.261)	w2d	(0.423)	dst	(0.547)
4	mlrt	(0.009)	w1d	(0.080)	w1d	(0.229)	w1d	(0.400)	w1d	(0.537)
5	w2d	(0.009)	mlrt	(0.072)	mlrt	(0.216)	mlrt	(0.388)	mlrt	(0.531)
6	adt	(0.009)	adt	(0.066)	adt	(0.208)	adt	(0.378)	adt	(0.526)
7	dst	(0.009)	elr	(0.049)	elr2	(0.158)	elr2	(0.324)	elr2	(0.493)
8	elr	(0.008)	elr2	(0.047)	elr	(0.154)	elr	(0.314)	elr	(0.488)
9	dtest	(0.006)	dtest	(0.016)	dtest	(0.056)	dtest	(0.175)	dtest	(0.392)
$p = 0.1$										
		$\theta = 1$		$\theta = 5$		$\theta = 10$		$\theta = 20$		$\theta = 50$
1	adds	(0.011)	w1d	(0.118)	mlrt	(0.316)	mlrt	(0.596)	elr2	(0.847)
2	dst	(0.011)	adds	(0.116)	w1d	(0.313)	adt	(0.563)	mlrt	(0.829)
3	adt	(0.011)	w2d	(0.114)	adt	(0.290)	elr2	(0.562)	adt	(0.823)
4	elr	(0.010)	mlrt	(0.112)	adds	(0.286)	w1d	(0.558)	elr	(0.811)
5	mlrt	(0.010)	dst	(0.103)	elr2	(0.264)	adds	(0.536)	adds	(0.795)
6	elr2	(0.010)	adt	(0.100)	w2d	(0.261)	elr	(0.507)	w1d	(0.756)
7	w2d	(0.009)	elr2	(0.082)	elr	(0.232)	w2d	(0.417)	dtest	(0.738)
8	w1d	(0.009)	elr	(0.080)	dst	(0.204)	dtest	(0.381)	w2d	(0.540)
9	dtest	(0.007)	dtest	(0.035)	dtest	(0.132)	dst	(0.304)	dst	(0.404)
$p = 0.9$										
		$\theta = 1$		$\theta = 10$		$\theta = 20$		$\theta = 50$		$\theta = 100$
1	mlrt	(0.011)	mlrt	(0.044)	elr2	(0.066)	elr2	(0.122)	elr	(0.184)
2	adds	(0.010)	elr2	(0.042)	elr	(0.063)	elr	(0.118)	elr2	(0.178)
3	elr2	(0.010)	elr	(0.036)	mlrt	(0.063)	mlrt	(0.104)	mlrt	(0.136)
4	adt	(0.010)	adt	(0.036)	adt	(0.053)	dtest	(0.096)	dtest	(0.134)
5	w1d	(0.010)	w1d	(0.033)	adds	(0.047)	adt	(0.094)	adt	(0.125)
6	elr	(0.010)	adds	(0.032)	w1d	(0.045)	adds	(0.082)	adds	(0.108)
7	dst	(0.010)	dtest	(0.029)	dtest	(0.044)	w1d	(0.062)	w1d	(0.062)
8	w2d	(0.009)	w2d	(0.023)	dst	(0.028)	w2d	(0.030)	dst	(0.030)
9	dtest	(0.008)	dst	(0.020)	w2d	(0.028)	dst	(0.029)	w2d	(0.028)

Table 4: Simulated power for different tests for 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)	elr	(0.024)	adt	(0.093)	adt	(0.461)	adt	(0.915)
7	elr	(0.010)	dtest	(0.021)	elr	(0.087)	elr	(0.367)	elr	(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)
$p = 0.5$										
		$\theta = 1$		$\theta = 2$		$\theta = 3$		$\theta = 5$		$\theta = 7$
1	dtest	(0.013)	mlrt	(0.098)	mlrt	(0.431)	mlrt	(0.922)	mlrt	(0.993)
2	adds	(0.011)	dst	(0.093)	w2d	(0.417)	w1d	(0.908)	w1d	(0.991)
3	elr	(0.010)	w2d	(0.093)	w1d	(0.371)	w2d	(0.907)	adt	(0.990)
4	adt	(0.010)	adds	(0.084)	adds	(0.341)	adt	(0.869)	adds	(0.987)
5	mlrt	(0.010)	w1d	(0.074)	dst	(0.328)	adds	(0.869)	w2d	(0.987)
6	dst	(0.009)	dtest	(0.059)	adt	(0.290)	dtest	(0.851)	dtest	(0.987)
7	elr2	(0.009)	adt	(0.051)	dtest	(0.276)	elr2	(0.833)	elr2	(0.986)
8	w1d	(0.008)	elr	(0.051)	elr2	(0.253)	elr	(0.785)	elr	(0.973)
9	w2d	(0.008)	elr2	(0.050)	elr	(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)	elr	(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)	elr	(0.230)	adt	(0.474)
4	elr2	(0.010)	elr	(0.018)	w1d	(0.088)	mlrt	(0.193)	dtest	(0.466)
5	mlrt	(0.010)	dst	(0.018)	elr	(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	elr	(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 5: Simulated power for different tests for 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)	elr	(0.022)	adt	(0.094)	adt	(0.421)	adt	(0.796)
8	elr	(0.010)	elr2	(0.022)	elr	(0.093)	elr	(0.331)	elr	(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)	elr	(0.015)	adds	(0.095)	adds	(0.676)	adds	(0.986)
7	elr	(0.009)	elr2	(0.014)	elr	(0.068)	adt	(0.563)	adt	(0.978)
8	mlrt	(0.009)	adds	(0.013)	elr2	(0.061)	elr	(0.464)	elr	(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)	elr	(0.996)
2	w2d	(0.012)	w2d	(0.079)	w2d	(0.268)	elr	(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	elr	(0.010)	dst	(0.057)	elr	(0.231)	w1d	(0.599)	w1d	(0.862)
6	elr2	(0.010)	elr	(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)

- 1 Balakrishnan, N., Basu, A. P., 1996. The Exponential Distribution: Theory, Methods, and
2 Applications. Gordon and Breach, New York.
- 3
4
- 5 Charnigo, R., Sun, J., 2004. Testing homogeneity in a mixture distribution via the l^2 distance
6 between competing models. Journal of the American Statistical Association 99, 488 – 498.
- 7
8
- 9
10 Chen, H., Chen, J., Kalbfleisch, J. D., 2001. A modified likelihood ratio test for homogeneity
11 in finite mixture models. Journal of Royal Statistical Society, Series B 63, 19 – 29.
- 12
13
- 14
15 Choi, S., 1979. Two-sample tests for compound distributions for homogeneity of mixing pro-
16 portions. Technometrics 21 (3), 361–365.
- 17
18
- 19 Ciuperca, G., 2002. Likelihood ratio statistic for exponential mixtures. Annals of the Institute
20 of Statistical Mathematics 54 (3), 585–594.
- 21
22
- 23
24 Efimova, T., Leskin, V., Ososkov, G., Tolstov, K., Chernov, N., 1989. Expansion of transverse
25 momenta in inelastic collisions of particles into rayleigh distributions. JINR Rapid Commu-
26 nications 3[36].
- 27
28
- 29
30 Garel, B., 2007. Recent asymptotic results in testing for mixtures. Computational Statistics
31 and Data Analysis 51, 5295–5304.
- 32
33
- 34
35 Groeneboom, P., Oosterhoff, J., 1977. Bahadur efficiency and probabilities of large deviations.
36 Statist. Neerlandica 31, 1–24.
- 37
38
- 39
40 Hartigan, J., 1985. A failure of likelihood ratio asymptotics for normal mixtures. In: proc.
41 Berkeley conference in honor of Jerzy Neyman and Jack Kiefer. Wasworth Advanced books,
42 Monterey, CA.
- 43
44
- 45
46 Hazan, A., Landsman, Z., Makov, U., 2003. Robustness via a mixture of exponential power
47 distributions. Computational Statistics and Data Analysis 42, 111–121.
- 48
49
- 50
51 Henze, N., Meintanis, S., 2005. Recent and classical tests for exponentiality: a partial review
52 with comparisons. Metrika 61, 29–45.
- 53
54
55
56
57
58

- 1 Höning, V., Stehlík, M., Danielová, V., Daniel, M., Švec, P., Grubhoffer, L., 2010. Tick host-
2 seeking activity and tick-borne encephalitis incidence: regression and homogeneity. Journal
3 of Applied Mathematics, Statistics and Informatics 6 (1), in press.
4
5
6
7 Lehmann, E., 1964. Testing Statistical Hypotheses. John Wiley & Sons, New York.
8
9
10 Li, P., Chen, J., Marriott, P., 2009. Non-finite fisher information and homogeneity: an em
11 approach. Biometrika 96 (2), 411–426.
12
13
14 Lindsay, B., 1995. Mixture models: Theory, geometry and applications. Institute of Mathemat-
15 ical Statistics, Hayward, Cal.
16
17
18 Liu, X., Pasarica, C., Shao, Y., 2003. Testing homogeneity in gamma mixture models. Scand.
19 J. Statist. 30, 227–239.
20
21
22
23 Manoukian, E., 1986. Modern concepts and theorems of mathematical statistics. Springer-
24 Verlag:Springer Series in Statistics, New York.
25
26
27
28 McLachlan, G., 1995. Mixtures - models and applications. the exponential distribution. In:
29 proc. Berkeley conference in honor of Jerzy Neyman and Jack Kiefer. Gordon & Breach,
30 Amsterdam, pp. 307–323.
31
32
33
34 Meintanis, S., 2007. Test for exponentiality against weibull and gamma decreasing hazard rate
35 alternatives. KYBERNETIKA 43 (3), 307–314.
36
37
38
39 Mosler, K., Haferkamp, L., 2007. Size and power of recent tests for homogeneity in exponential
40 mixtures. Communications in Statistics - Simulation and Computation 36, 493–504.
41
42
43
44 Mosler, K., Haferkamp, L., 2009. A comparison of recent procedures in weibull mixture testing.
45 In: Advances in Data Analysis. Birkhäuser, Boston.
46
47
48
49 Mosler, K., Scheicher, C., 2008. Homogeneity testing in a weibull mixture model. Statistical
50 Papers 49, 315–332.
51
52
53 Mosler, K., Seidel, W., 2001. Testing for homogeneity in an exponential mixture model. Aus-
54 tralian and New Zealand Journal of Statistics 43, 231–247.
55
56
57
58
59
60

- 1 Randles, R. H., 1982. On the asymptotic normality of statistics with estimated parameters.
2
3 Ann. Statist. 10, 462–474.
4
- 5 Rublík, F., 1989a. On optimality of the LR tests in the sense of exact slopes, part 2, application
6
7 to individual distributions. Kybernetika 25, 117–135.
8
- 9
10 Rublík, F., 1989b. On optimality of the LR tests in the sense of exact slopes, Part 1, general
11
12 case. Kybernetika 25, 13–25.
13
- 14 Seidel, W., Mosler, K., Alker, M., 2000. A cautionary note on likelihood ratio tests in mixture
15
16 models. Annals of the Institute of Statistical Mathematics 52, 481–487.
17
- 18
19 Stehlík, M., 2003. Distributions of exact tests in the exponential family. Metrika 57, 145–164.
20
- 21
22 Stehlík, M., 2006. Exact likelihood ratio scale and homogeneity testing of some loss processes.
23
24 Statistics and Probability Letters 76, 19–26.
25
- 26
27 Stehlík, M., 2008. Homogeneity and scale testing of generalized gamma distribution. Reliability
28
29 Engineering & System Safety 93, 1809–1813.
30
- 31
32 Stehlík, M., Ososkov, G., 2003. Efficient testing of the homogeneity, scale parameters and
33
34 number of components in the rayleigh mixture. JINR Rapid Communications E-11-2003-
35
36 116.
37
- 38
39 Susko, E., 2003. Weighted tests of homogeneity for testing the number of components in a
40
41 mixture. Computational Statistics and Data Analysis 41, 367–378.
42
- 43
44 Tchirina, A., 2005. Large deviations for a class of scale-free statistics under the gamma distri-
45
46 bution. Journal of Mathematical Sciences 128(1), 2640–2655.
47
- 48
49 Wilks, S., 1938. The large sample distribution of the likelihood ratio for testing composition
50
51 hypotheses. Ann. Math. Statist. 9, 60–62.
52
53
54
55
56
57
58
59
60