Nonlinear Studies - www. nonlinearstudies.com
MESA - www.journalmesa.com

Preprint submitted to Nonlinear Studies / MESA

I-divergence based testing with applications to reliability and physics

Milan Stehlík[1,⋆], Jozef Kiseľák[2] and Gennadij A. Ososkov[3]

[1] Department of Applied Statistics, Altenberger Straße 69, Johannes Kepler University in Linz, A-4040 Linz, Austria. [2] Institute of Mathematics, Faculty of Science, P.J.Šafárik University in Košice.
[3] Joint Institute for Nuclear Research, Dubna, Moscow region 141980, Russia
⋆*Corresponding Author*. Email: Milan.Stehlik@jku.at

Abstract. We introduce a test based on the *I*-divergence from the observed vector to the canonical parameter (see [1]). Such a test is employed for testing for homogeneity and the value of the scale parameter in the exponential family. We illustrate the applicability of such a test for real data on light indicators for aeroplanes. We also show how we can extend such a construction outside the exponential family. This will be shown on the statistical problem to expand the experimental distribution of transverse momenta into Rayleigh distribution in photoemulsion experiments. By such an expansion we will construct a high-efficient testing procedure for testing the hypothesis of the homogeneity.

1 Introduction

We introduce a test based on the I-divergence from the observed vector to the canonical parameter $I_N(y,\gamma)$ (see [1]), which represents the distance (based on N observations) between the observed vector y and the hypothesized model represented by its parameter γ.

Such a test is employed on testing for homogeneity and the value of the scale parameter in the exponential family. We illustrate the applicability of such a test for real data on light indicators for aeroplanes. We also show how we can extend such a construction outside the exponential family. This will be illustrated on the statistical problem to expand the experimental distribution of transverse momenta into Rayleigh distribution in photoemulsion experiments.

The paper is organized as follows. In Section 2 we introduce the test based on $I_N(y,\gamma)$ and apply it to a real data example on airplane indicator light operating times. A simulation study in R reveals the properties of such a test. Here we also discuss the numerical complexity of these simulations. Some

2010 **Mathematics Subject Classification**
Keywords: information divergence, deconvolution, decomposition, statistical inference, homogeneity testing, number of mixture components testing, scale testing, Rayleigh mixture, subpopulation model, exact distribution, asymptotic efficiency, Lambert W function

general notes on exact distributions are also given. In section 3 open problems, mainly related to the decomposition of I-divergence are mentioned. Section 4 extends testing based on the decomposition of I-divergence outside of the exponential family. This will be illustrated on the statistical problem to expand the experimental distribution of transverse momenta into Rayleigh distribution in photoemulsion experiments. Section 5 concludes. In the Appendix (Section 6) the properties of the Lambert W (LW) function used in the paper are shortly recalled.

2 Testing for homogeneity and scale of airplane light indicators

In many practical cases, the practitioner (e.g.reliability engineer) is interested only in testing of particular life time (i.e. scale hypothesis under uncertainty about homogeneity in the sample). One remedy in such a setup is to test a hypothesis based on an I-divergence $I_N(y,\gamma_0)$ directly, i.e. to statistically measure the deviation of the observed vector y from the hypothesized canonical parameter γ_0. This will be illustrated on the test for lifetime of airplane light indicators in the next section. Here we consider as a motivating example a real data set (see Table 1) of airplane indicator light operating times from Reliability Analysis Center (RAC) database (see [2]). Here $N = 6$ is not the number of observations since the times provided are the cumulative times. According to the number of failures the number of aggregated observations is 38.

Failures	T_j	Cumulative operating time (hours)
2	T_1	51 000
9	T_2	194 900
8	T_3	45 300
8	T_4	112 400
6	T_5	104 000
5	T_6	44 800

Table 1 Airplane indicator light reliability data

2.1 Testing by the exact LR test for the scale parameter

If the homogeneity of the scale parameters is statistically significant, we may use a directed test for scale parameter (with simple null, and composite alternative hypothesis, typically). In [2] we can find the MLE of shape parameter ($v = 0.7$) and scale parameter ($\gamma = 0.0000484$) of the gamma distributed individual times-to-failure of the data in Table 1.

We have $\omega = 38 \times 0.7 = 26.6$ and $\sum_{j=1}^{6} T_j = 552\,400$.

Let us consider the testing problem

$$H_0 : \gamma = 0.00003207 \text{ versus } H_1 : \gamma \neq 0.00003207 \tag{2.1}$$

at the level of significance 0.05.

In particular a reliability practitioner could be interested in conducting the hypothesis (2.1) test, to see whether the field reliability has significantly changed from its current level.

The critical value $c_{0.05}$ of the exact LR test of the hypothesis (2.1) is $c_{0.05} = 3.86550298$.

We have

$$-2\ln\Lambda = 3.855303 < c_{0.05},$$

where Λ is the likelihood function. Therefore the null hypothesis is accepted at the level 0.05.

The power function $p(\gamma, 0.05)$ of the LR test of the hypothesis (2.1) has the form

$$1 - F^{\Gamma}_{26.6}\left(-\frac{26.6\gamma}{0.00003207}W_{-1}\left(-e^{-\frac{57.06550298}{53.2}}\right)\right) +$$

$$+F^{\Gamma}_{26.6}\left(-\frac{26.6\gamma}{0.00003207}W_{0}\left(-e^{-\frac{57.06550298}{53.2}}\right)\right)$$

e.g. for $\gamma \in (0.00001, 0.00007)$, where $F^{\Gamma}_{26.6}$ denotes cumulative distribution function of Gamma random variable with shape parameter $\omega = 26.6$ and scale parameter 1, and W_{-1}, W_0 are two branches of LW function (see also Appendix 6.1).

2.2 Testing by the exact test given by *I*-divergence $I_N(y, \gamma_0)$

Alternatively to the method in the previous section, we may use directly a test based on $I_N(y, \gamma_0)$, where γ_0 is the value of the scale parameter under the null. Consider a statistical model with N independent observations $y_1, \ldots, y_N$ which are distributed according to gamma densities (see [3])

$$f(y_i|\vartheta) = \begin{cases} \gamma_i(\vartheta)^{v_i}\frac{y_i^{v_i-1}}{\Gamma(v_i)}\exp(-\gamma_i(\vartheta)y_i), & \text{for } y_i > 0, \\ 0, & \text{for } y_i \leq 0. \end{cases} \tag{2.2}$$

Here $\gamma := (\gamma_1, \ldots, \gamma_N)$ is the vector of unknown scale parameters, which are the parameters of interest and $v = (v_1, ..., v_N)$ is the vector of known shape parameters. The parameter space Θ is an open subset of $\mathbf{R}^p$, $\gamma_i \in C^2(\Theta)$ and the matrix of first order derivatives of the mapping $\gamma := (\gamma_1, \ldots, \gamma_N)$ has full rank on Θ. This model is motivated e.g. by a situation when we observe time intervals between $(N+1)$ successive random events in a Poisson process. In this case the parameters γ_i are equal to the (usually parametrized) intensity γ.

Another interesting problem is a test for the proportion of the exponential distribution which can be used for constructing a statistical quality control sampling plan (see [4]).

By the use of covering property we can define the *I*-divergence of the observed vector y in the sense of [1] as

$$I_N(y, \gamma) := I(\hat{\gamma}_y, \gamma) = -\sum_{i=1}^{N}\{v_i - v_i\ln(v_i)\} + \sum_{i=1}^{N}\{y_i\gamma_i - v_i\ln(y_i\gamma_i)\}, \tag{2.3}$$

where $\hat{\gamma}_y$ denotes the maximum likelihood estimator of canonical parameter γ under observed vector y.

2.3 Simulation Study

In this section we provide simulation experiments with both real and simulated data to explore the properties of the *I*-divergence based test defined by the rejection region $I_N(y, \gamma) > C_\alpha$.

We have used the R environment ([5]) using function `rgama()`. Number of simulations were chosen from 10^4 to $5 \cdot 10^5$ depending on the parameters. In Figures 1 and 2 p-values of the real data for γ with fixed v and for v with fixed γ respectively are shown. Simulations were performed with critical constant $I_N^{\text{real}} = 8.13434$ computed from real data of cumulative operating time (hours) using the relation (2.3). In the pictures 3-6 critical values parameter v are shown. They were generated at the level of significance $\alpha = 0.01$ with constant γ and $N = 6, 25, 100, 1000$ number of observations (for Gamma distribution) respectively. In the next four pictures 7-10 are similar simulation results shown but at the level of significance $\alpha = 0.05$. Figures 11-14 and 15-18 shows similar simulations for constant $v = 0.7$ at the significance levels $\alpha = 0.05, 0.01$ respectively (again for $N = 6, 25, 100, 1000$). Finally, figures 19 and 20 show p-values for data generated from exponential distribution with $N = 100$ and constant $v = 1$ and $\gamma = 1$ respectively.

Example of simulated critical constants:

We simulated (2.3) for $y_i \sim Gamma$ with shape parameter 0.7 and scale parameter from the null hypothesis (2.1), i.e. $\gamma = 0.00003207$.

The following Table 2 displays obtained critical constants C_α for $\alpha = 0.05, 0.01$, for sample sizes $N = 6$ (because of data from Table 1), and $N = 25, N = 100, N = 1000$.

N	$c_{0.05}$	$c_{0.01}$
6	7.467411	9.707604
25	22.55542	26.15369
100	74.79102	81.23323
1000	648.1965	668.9786

Table 2 Critical constants C_α

The value of statistics for data given in Table 1 is $I_N^{\text{real}} = 8.13434$, having p-value=0.034, for number of simulations $nsim = 10000$.

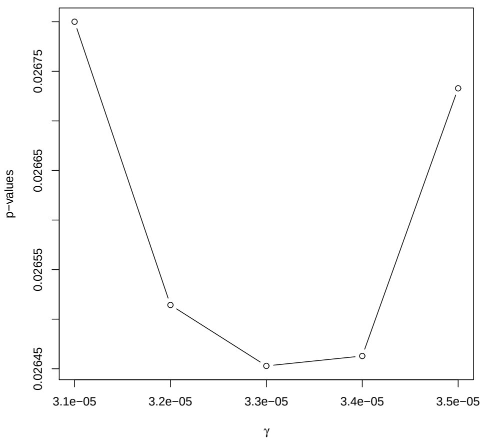

Fig. 1 P-values with parameters $v = 0.7$, $N = 6$.

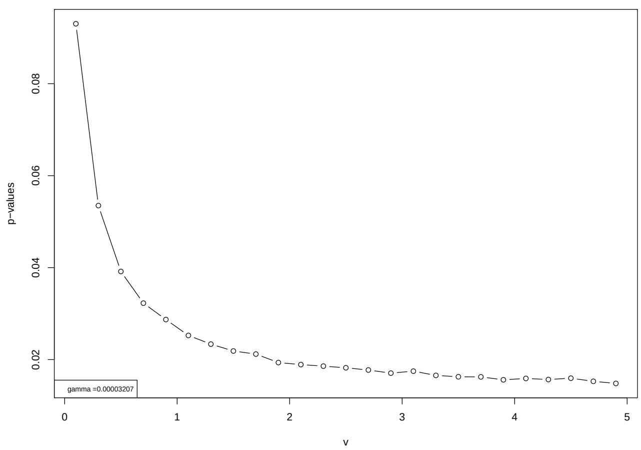

Fig. 2 P-values with parameters $\gamma = 3.207 \times 10^{-5}$, $N = 6$.

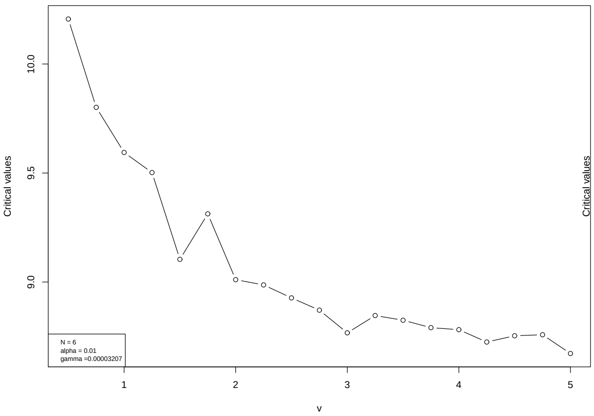

Fig. 3 Critical values with parameters $\gamma = 3.207 \times 10^{-5}$, $\alpha = 0.01$, $N = 6$.

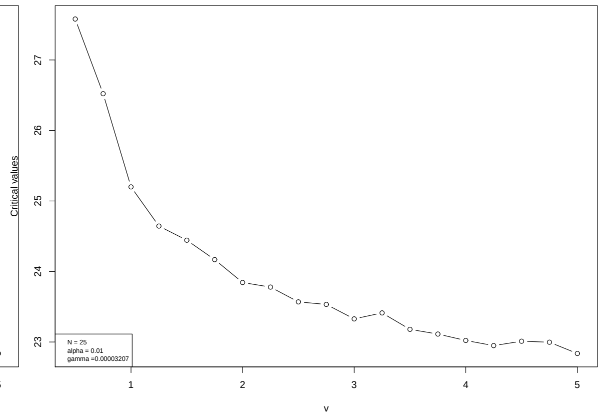

Fig. 4 Critical values with parameters $\gamma = 3.207 \times 10^{-5}$, $\alpha = 0.01$, $N = 25$.

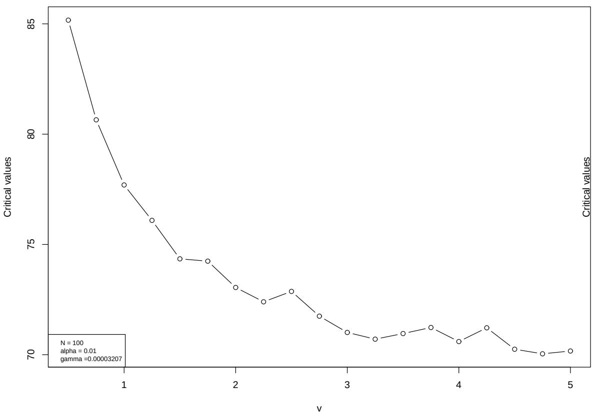

Fig. 5 Critical values with parameters $\gamma = 3.207 \times 10^{-5}$, $\alpha = 0.01$, $N = 100$.

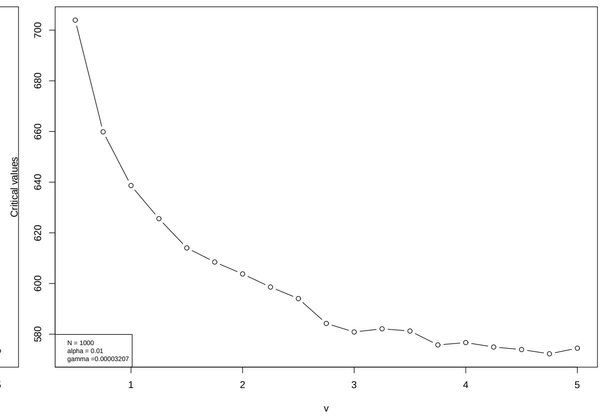

Fig. 6 Critical values with parameters $\gamma = 3.207 \times 10^{-5}$, $\alpha = 0.01$, $N = 1000$.

2.4 Numerical problems of simulations

Time complexity depends on the number of simulations, the number N of generated realizations of gamma distribution (not for p-values) and the number of discretization points of parameter interval.

For illustration:

Procedure computing p-values:
one p-value for $N = 6$ and nsim= 10^5 lasts 13,5 s
one p-value for $N = 6$ and nsim= 10^6 lasts 164,3 s

Procedure computing critical values:
one p-value for $N = 6$ and nsim= 10^5 lasts 17 s
one p-value for $N = 25$ and nsim= 10^4 lasts 5,5 s
one p-value for $N = 100$ and nsim= 10^4 lasts 91,4 s

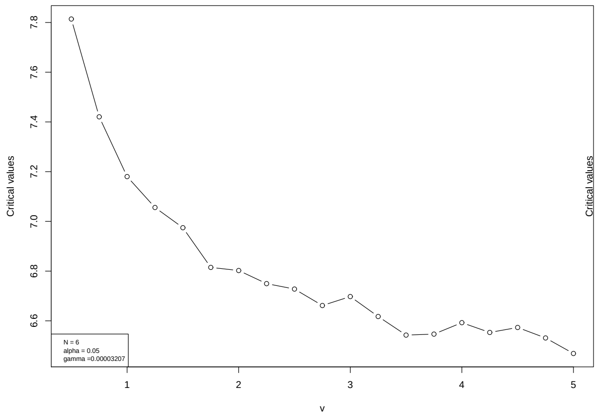

Fig. 7 Critical values with parameters $\gamma = 3.207 \times 10^{-5}$, $\alpha = 0.05$, $N = 6$.

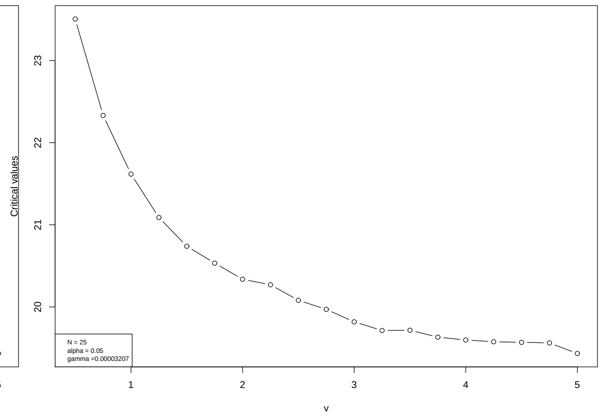

Fig. 8 Critical values with parameters $\gamma = 3.207 \times 10^{-5}$, $\alpha = 0.05$, $N = 25$.

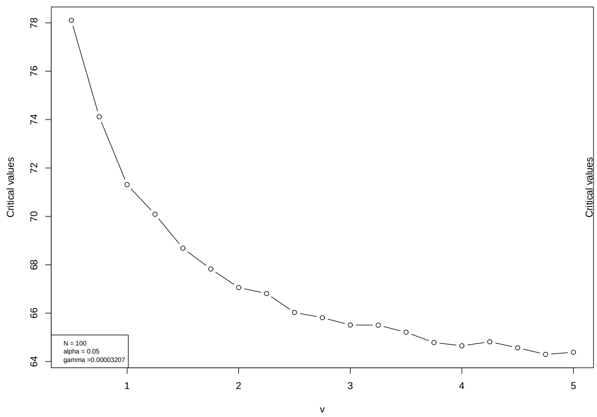

Fig. 9 Critical values with parameters $\gamma = 3.207 \times 10^{-5}$, $\alpha = 0.05$, $N = 100$.

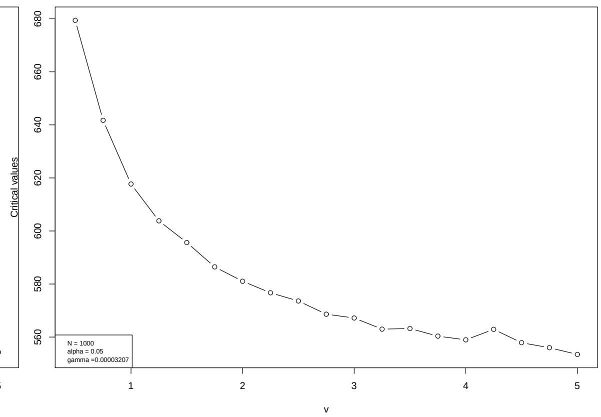

Fig. 10 Critical values with parameters $\gamma = 3.207 \times 10^{-5}$, $\alpha = 0.05$, $N = 1000$.

These times should be multiplied by the number of discrete points of the parameters v and γ.

2.5 General notes

To derive the distribution of the test based on the exact I-divergence we can use geometric measure integration. For the two-dimensional case, [3] has derived (see Theorems 1 and 2 therein) that $I_2(y, \delta(1,1)) = R_2 + S_2$ where R_2, S_2 are independent. Here $\delta(1,1)$ denotes multiplication of true canonical parameter vector $(1,1)$ by scalar δ, which allows scaling of true values of canonical parameters.

The c.d.f. of the random variable R_N has the form

$$F_N(\rho) = \begin{cases} \mathcal{F}_N(-\frac{N}{\delta}\, W_{-1}(-\exp(-1-\frac{\rho}{N}))) - \mathcal{F}_N(-\frac{N}{\delta}\, W_0(-\exp(-1-\frac{\rho}{N}))), & \rho > 0, \\ 0, & \rho \leq 0 \end{cases}$$

and the density of R_N has the form

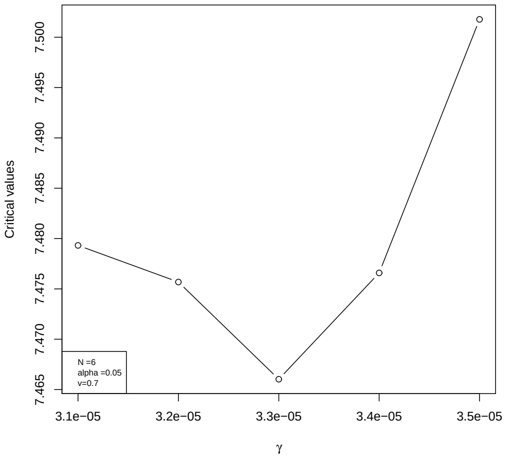

Fig. 11 Critical values with parameters $v = 0.7,\ \alpha = 0.05,\ N = 6$.

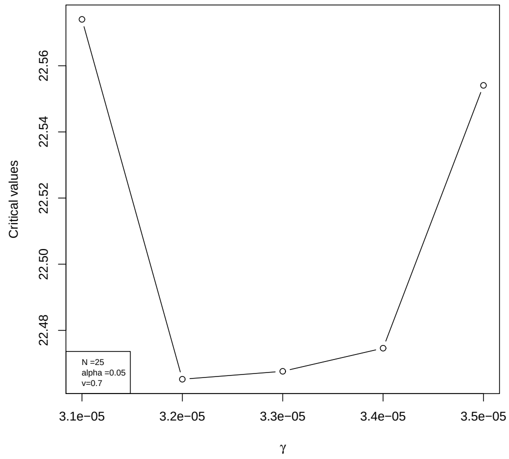

Fig. 12 Critical values with parameters $v = 0.7,\ \alpha = 0.05,\ N = 25$.

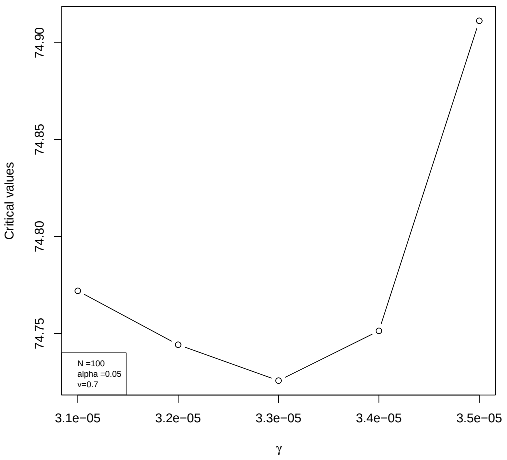

Fig. 13 Critical values with parameters $v = 0.7,\ \alpha = 0.05,\ N = 100$.

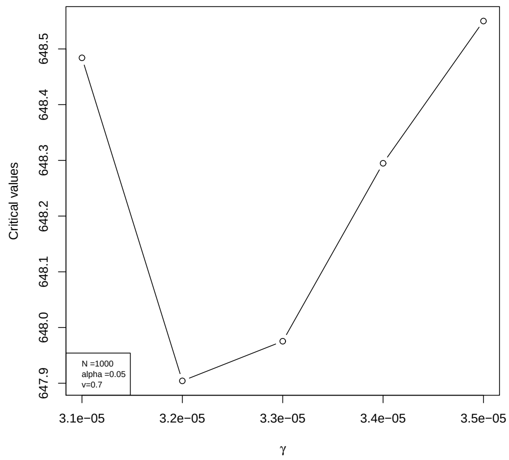

Fig. 14 Critical values with parameters $v = 0.7,\ \alpha = 0.05,\ N = 1000$.

$$f_N(\rho) = \begin{cases} h(N,1,\rho,\delta^{-1}) - h(N,0,\rho,\delta^{-1}), & \text{for } \rho > 0, \\ 0, & \text{for } \rho \leq 0. \end{cases}$$

Here $\mathcal{F}_N$ is the c.d.f. of the $\Gamma(N,1)$-distribution and for $\tau, r, s > 0;\ k \in \mathbf{Z}$ we define

$$h(N,k,r,s) = \frac{(-N)^{N-1} s^N}{\Gamma(N)} \frac{\{W_{-k}(-\exp(-1-\frac{r}{N}))\}^N}{1 + W_{-k}(-\exp(-1-\frac{r}{N}))} \exp\left(NsW_{-k}\left(\exp\left(-1-\frac{r}{N}\right)\right)\right),$$

where W_0, W_{-1} are two real-valued branches of Lambert-W function.

Under the null hypothesis of homogeneity, cdf of S_2 has the form (see [3]):

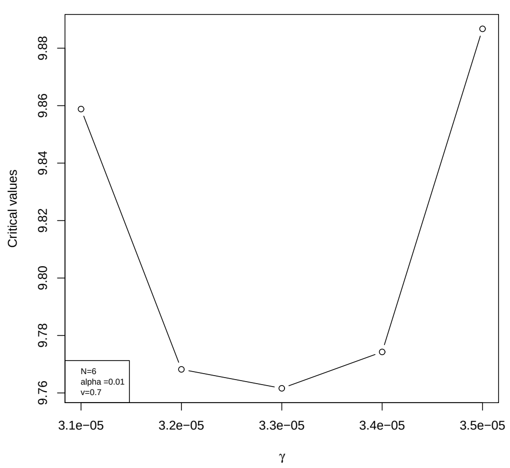

Fig. 15 Critical values with parameters $v = 0.7,\ \alpha = 0.01,\ N = 6$.

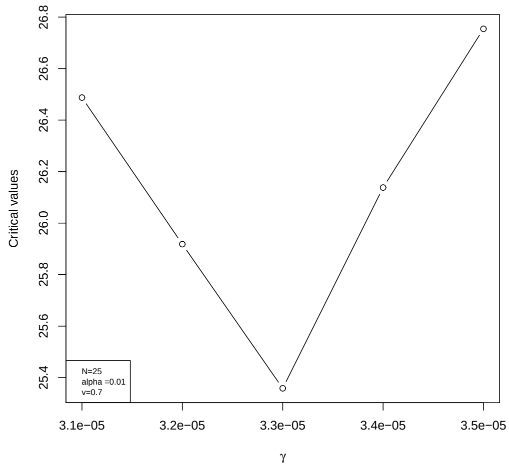

Fig. 16 Critical values with parameters $v = 0.7,\ \alpha = 0.01,\ N = 25$.

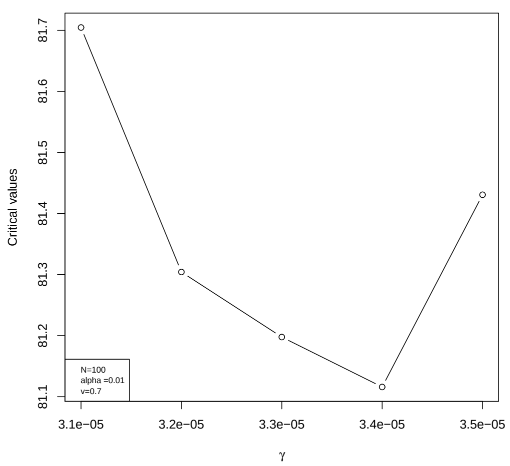

Fig. 17 Critical values with parameters $v = 0.7,\ \alpha = 0.01,\ N = 100$.

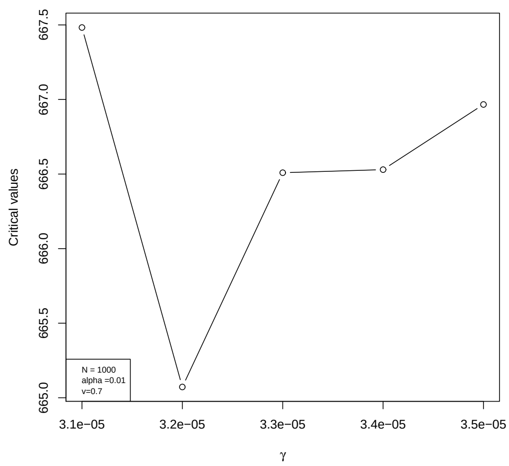

Fig. 18 Critical values with parameters $v = 0.7,\ \alpha = 0.01,\ N = 1000$.

$$F_2(x) = \begin{cases} \sqrt{1-\exp(-x)}, & \text{for } x > 0, \\ 0, & \text{for } x \leq 0 \end{cases}$$

Thus, the sum $R_2 + S_2$ has a density which is a convolution of both densities. Generally, we are working with the components S_N, R_N in a separate manner, since these components correspond to the popular likelihood ratio tests of the homogeneity

$$H_0 : \gamma_1 = \gamma_2 = ... = \gamma_N \text{ versus } H_1 := non\ \ H_0 \tag{2.4}$$

and scale hypothesis

$$H_0 : \gamma = \gamma_0 \text{ versus } H_1 : \gamma \neq \gamma_0 \tag{2.5}$$

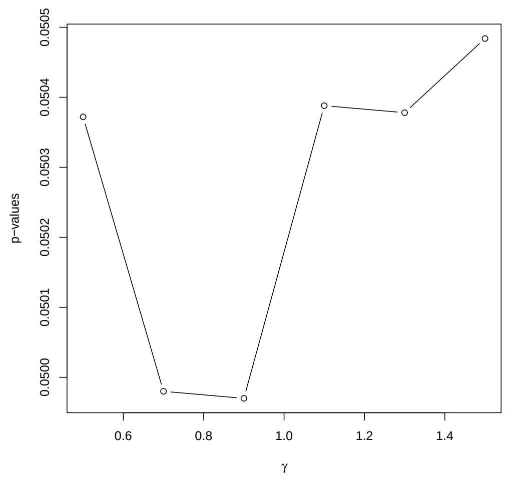

Fig. 19 P-values with parameters $v = 1$, $N = 100$.

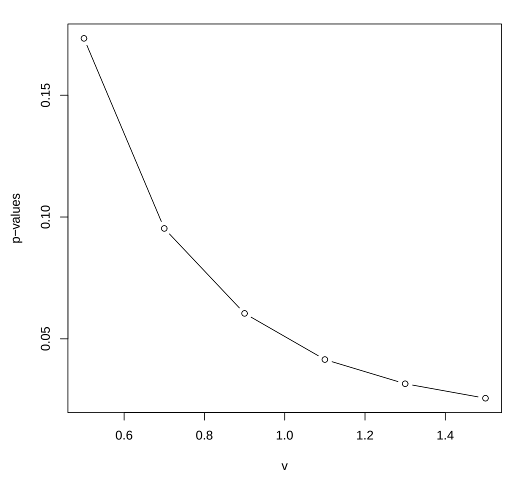

Fig. 20 P-values with parameters $\gamma = 1$, $N = 100$

when the sample is driven from the gamma distribution. We may generalize the classical gamma distribution to the so-called generalized gamma family with density of the form

$$f(y_i|\vartheta) = \frac{\alpha}{\sigma\Gamma\left(\frac{1+\beta}{\alpha}\right)} \left(\frac{y_i}{\sigma}\right)^{\beta} \exp\left(-\left(\frac{y_i}{\sigma}\right)^{\alpha}\right),$$

for $y_i > 0$, and $\vartheta = (\alpha, \beta, \sigma)$. The generalized gamma distribution has many applications in life sciences, reliability theory, engineering and physics. The generalized gamma distribution is one of the most studied probability density functions of statistics since many of the important nondiscrete density functions can be derived from it. For example, $f(y|(2,0,\sqrt{2\sigma}))$ is the one-sided normal distribution, and $f(y|(1,n/2-1,2))$ is the χ^2_n-distribution. In the special case of $\beta = \alpha - 1$ the gamma distribution is called a Weibull distribution and in case of $\alpha = 1$ we obtain the Gamma distribution. While not as frequently used for modeling life data as the previous distributions, the generalized gamma distribution does have the ability to mimic the attributes of other distributions such as the Weibull or lognormal, based on the values of the distribution's parameters.

To the best of our knowledge, the exact likelihood ratio (LR) test for the scale and homogeneity in the complete sample from the gamma family has been derived in [3]. In [6] the exact likelihood ratio test for the scale and homogeneity in the complete sample from the Weibull family is derived. In [7] the exact likelihood ratio test for the scale and homogeneity in the complete sample from the generalized gamma family is derived. The exact likelihood ratio test for the scale parameter in the Type I, Type II and progressively Type II censored sample is derived in [8]. The approach for the exact likelihood ratio testing for the scale and homogeneity with the missing time to failure exponential data is given by [9]. These tests have been shown to be optimal in the sense of Bahadur (see [10], [11] and [3]). The exact LR test for the scale has asymptotically a χ^2 distribution (this is the result of Samuel Wilks, see [12]).

3 Open Problems

It is clear that efficient or somehow optimal statistical decisions are related to the information divergences or their decompositions. Several open problems remain.

In the case of the I-divergence and the gamma family, we have an interesting decomposition of I-divergence from the observed vector to the canonical hypothesized parameter to the LR statistics of scale and homogeneity discussed above. In [13] a generalized family of measures of divergence is investigated and applied successfully in statistical inference. Similar deconvolution ideas will be of further interest also for such families of divergences.

The second open problem, discussed in more detail in [14], is the relation of the ϕ-*divergences* and *statistical information* in a dimension higher than 2. The case $n = 2$ was thoroughly studied in [15].

One can consider also the decomposition of the expected Kullback divergence corresponding to the expected discrepancy with Akaike's information criterion ([16]) as presented for the state-space framework in [17]. This has a nice statistical interpretation in terms of the *expected optimism.* Further study in this direction will be of interest.

4 Example of photoemulsion experiment

This section is devoted to the statistical problem of expanding an experimental distribution of transverse momenta $P_\perp$ into a series of Rayleigh distributions and can be considered as a continuation of [18]. The physical background of this problem arises in the emulsion experiment studying the dynamics of inelastic collision of fast heavy particles as nuclei ^{22}Ne with the photoemulsion nuclei by momenta 4.1 A GeV/c. The spectrum of transverse momenta for inclusive experiment can bear the quite important information about the generation process of secondary particles, whether this process is direct or is going through some intermediate stages. As it is known (see, for example, [18]), transverse momenta are distributed according to the Rayleigh law. However depending on the collision model (one of more than one channels of the particle generation) the $P_\perp$ distribution can be described by just one Rayleigh distribution or by a series

$$f(y;P_\perp) = \sum_{i=1}^{k} a_i \frac{y}{\sigma_i^2} \exp\left(-\frac{y^2}{2\sigma_i^2}\right), y > 0, \sum a_i = 1$$

with some unknown k, σ_i and a_i. The formulation of mathematical problem is complicated due to experimental restrictions caused by different conditions of registering secondary particles depending on the emanating angle θ of those particles in respect to the collision axis. That was taken into account in [18] by inventing corresponding statistical weights of measured $P_\perp$ depending on θ and allowed to elaborate a method of expanding the experimental $P_\perp$-distribution into one or the mixture of two Rayleigh distributions.

However the generalization of the Rao-Smirnov ω^2 test proposed in [18] to choose the hypothesis about the expansion type (one or the mixture of two Rayleigh distributions) was not proven to be optimal. Therefore at the present section we focus on constructing of the high efficient testing procedure of the homogeneity hypothesis with general and mixture alternatives.

The likelihood-ratio decision procedure related to the hypothesis $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1 \setminus \Theta_0$ is based on 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 interest parameter and $L_y(\theta)$ is the likelihood of θ under the observed y. The maximum likelihood principle has been very successful in leading to satisfactory procedures in many specific problems (see [19] for details).

In Section 4.1 we approximate the mixture model by the subpopulation one. In Section 4.2 we relate tests from the exponential family to the test with Rayleigh distributions, outside of the exponential family. This relation enables us to use the I-divergence test also for Rayleigh distributions. In Section 4.3 we discuss the procedure for the LR testing of the hypotheses of the number of components m in the Rayleigh mixture for $m = 2$ and 3. In the Appendix we provide some properties of the Lambert W function.

4.1 Subpopulation model

There are physical reasons for considering $40 \leq N \leq 50$. In smaller samples we recommend the exact LR testing for the number of components m in the mixture. Such a procedure, however, leads to a rather laborious computation. Practical difficulties arise specially due to the likelihood frequently having multiple local extremes. In our approach we approximate the exact mixture model given by the density $\prod_{i=1}^{N} f(y_i|\sigma^2)$ where $f(y|\sigma^2)$ is the mixture density

$$f(y|\sigma^2) = \sum_{i=1}^{m} \pi_i \frac{y}{\sigma_i^2} \exp\left(-\frac{y^2}{2\sigma_i^2}\right), \ \pi_1 + \ldots + \pi_m = 1 \tag{4.1}$$

with the subpopulation model given by the density $\prod_{i=1}^{k_1} f(y_{n_1,i}|\sigma_1^2)\ldots\prod_{i=1}^{k_m} f(y_{n_m,i}|\sigma_m^2)$. Here observations $y_{n_j,1},\ldots,y_{n_j,k_j}$ belong to the j-th subpopulation, $k_1 + k_2 + \ldots + k_m = N$ and $\frac{k_j}{N}$ approximates the probability π_j of selecting an individual from subpopulation j. The component density $f(y|\sigma_j^2) = \frac{y}{\sigma_i^2} \exp\left(-\frac{y^2}{2\sigma_i^2}\right), y > 0$ is the conditional density of Y given that the observation is from the j-th subpopulation.

The subpopulation model is frequently used as motivation for the mixture density (see [20]) in large samples. Since the true classification of observations into subpopulations is unobserved, the density (4.1) is typically used for the observations.

4.2 Homogeneity testing

In this section we derive the exact distribution of the LR test for homogeneity of the Rayleigh distribution. We consider a statistical model with N independent observations $y_1,\ldots,y_N$ which are distributed according to Rayleigh densities

$$f(y_i|\sigma_i^2) = \begin{cases} \frac{y_i}{\sigma_i^2} \exp\left\{-\frac{y_i^2}{2\sigma_i^2}\right\}, & \text{for } y_i > 0, \\ 0, & \text{for } y_i \leq 0. \end{cases} \tag{4.2}$$

Here $\sigma^2 := (\sigma_1^2,\ldots,\sigma_N^2)$ is the vector of unknown scale parameters. Let us introduce the notation $X \sim R(\sigma^2)$ when X is distributed according to density (4.2) with the scale parameter σ^2 and $X \sim Exp(\gamma)$ when X is distributed according to the exponential density

$$f(x|\gamma) = \begin{cases} \gamma \exp\{-\gamma x\}, & \text{for } x > 0, \\ 0, & \text{for } x \leq 0 \end{cases}$$

with the scale parameter γ. Now let us construct the efficient test of the homogeneity in the model (4.2). The null hypothesis has the form

$$H_0:\ \sigma_1^2 = \ldots = \sigma_N^2. \tag{4.3}$$

The LR of the homogeneity test has the form

$$\lambda_N(y) = \frac{\max\limits_{\sigma_1^2=\ldots=\sigma_N^2} f(y,\sigma^2)}{\max\limits_{\sigma^2} f(y,\sigma^2)},$$

where $f(y,\sigma^2) = \prod_{i=1}^{N} f(y_i|\sigma_i^2)$. After the optimization we obtain that

$$\lambda_N(y) = \frac{N^N (y_1 \ldots y_N)^2}{(y_1^2 + \ldots + y_N^2)^N}. \tag{4.4}$$

If $X \sim R(\sigma^2)$ holds, then we have

$$\frac{X^2}{2\sigma^2} \sim Exp(1). \tag{4.5}$$

Relationship (4.5) is substantial for relating the homogeneity and scale testing in the exponential family (exp. distribution) and outside the exponential family (Rayleigh). Thus simulation results obtained in section 2 for the gamma distribution can be related to the analogous tests of the Rayleigh distribution. Under the homogeneity hypothesis, the distribution of the likelihood ratio (4.4) does not depend on the unknown parameter σ^2. Furthermore, due to (4.5) we have that $\lambda_N(y)$ has under H_0 the same distribution as the homogeneity LR statistics

$$\frac{N^N x_1 \ldots x_N}{(x_1 + \ldots + x_N)^N}$$

of the homogeneous exponential sample $x_1, \ldots, x_N$ (see [3]). Due to the monotonous transformation $g(x) = \sqrt[N]{x}$ of the likelihood ratio (4.4) we obtain the interesting statistics of the homogeneity,

$$\frac{\sqrt[N]{y_1^2 \ldots y_N^2}}{\frac{y_1^2 + \ldots + y_N^2}{N}}.$$

that is the ratio of the geometric and arithmetic mean of the squares of observations. The distribution of the LR test statistics $-\ln \lambda_N$ of the homogeneity under the null hypothesis is derived in [6].

Computation of the critical values of the test uses the fact that $\lambda_N(y)$ has under H_0 the same distribution as the homogeneity LR statistics

$$\frac{N^N x_1 \ldots x_N}{(x_1 + \ldots + x_N)^N}$$

where x_i are iid $Exp(1)$. For small dimensions we can compute the critical values from the exact c.d.f.s F_N of the test statistics $-\ln \lambda_N$. In [3] we can find, that in dimension 2 and 3 the c.d.f. has form

$$F_2(x) = \begin{cases} \sqrt{1-\exp(-x)}, & \text{for } x > 0, \\ 0, & \text{for } x \leq 0 \end{cases}$$

and

$$F_3(x) = \begin{cases} 2\int_{a(x)}^{b(x)} \frac{1}{s}\sqrt{s^2(1-s)^2 - \frac{4}{27}s\exp(-x)}ds, & \text{for } x > 0, \\ 0, & \text{for } x \leq 0 \end{cases}$$

where $0 < a(x) < b(x) < 1$ are solutions of the algebraic equation

$$t(1-t)^2 = \frac{4}{27}\exp(-x).$$

In high dimensions the c.d.f.s and densities are much more complicated and estimates of the critical values can be obtained by the simulation. These critical constants can be found in [21].

4.3 Efficient testing of the number of components in the Rayleigh mixture

In this section we discuss the efficient testing procedure of the number of components m in the Rayleigh mixture for $m = 2$ and 3. The case of the alternative hypothesis $H_1 : m = 2$ against the general alternative was thoroughly studied in [21], where also comparison to several commonly used tests for homogeneity has been conducted.

The case of the alternative $H_1 : m = 3$

In this section we consider the alternative of the form $H_1 : m = 3$. The hypothesis

$$H_0 : m = 1 \text{ versus } H_1 : m = 3 \tag{4.6}$$

in the mixture model (4.1) can be approximated due to the subpopulation model by the hypothesis

$$H_0 : \ \sigma_1^2 = ... = \sigma_n^2 \ \text{ versus } approx\ H_1 : \exists \text{ nonempty disjoint subsets } M_1, M_2, M_3 \tag{4.7}$$

$$\text{of the set}\{1, ..., n\}, \text{such that } \forall j \in M_1 : \sigma_j^2 = \sigma_1^2, \forall j \in M_2 : \sigma_j^2 = \sigma_2^2, \forall j \in M_3 : \sigma_j^2 = \sigma_3^2$$

where σ_1^2, σ_2^2 and σ_3^2 are different scale parameters.

We construct the LR test of the hypothesis (4.7) which approximates the hypothesis (4.6) . Let $y_1, ..., y_N$ are distributed according to Rayleigh densities. The LR of the test of the hypothesis (4.7) has the form

$$\lambda_N(y) = \frac{\max\limits_{\sigma_1^2 = ... = \sigma_N^2} f(y, \sigma^2)}{\max\limits_{approx\ H_1} f(y, \sigma^2)}.$$

To compute the denominator $\max\limits_{approx\ H_1} f(y, \sigma^2)$ we proceed as follows. Suppose that $\{y_{i_1}, .., y_{i_K}\}$, $0 < K < N-1$, are the observations from the Rayleigh distribution with the scale parameter σ_1^2, $\{y_{j_1}, .., y_{j_L}\}$, $0 < L < N-K$, are the observations from the Rayleigh distribution with the scale parameter σ_2^2 and the remaining observations are distributed according to the Rayleigh distribution with the scale parameter σ_3^2. For $0 < K < N-1$, $0 < L < N-K$ let $P(K, L)$ denote all disjoint pairs of K-subsets $\{i_1, ..., i_K\}$ and L-subsets $\{j_1, ..., j_L\}$ of the set $\{1, 2, ..., N\}$. Then the LR of the test of the hypotheses (4.7) has the form

$$\lambda_N(y) = \min_{0<K<N-1,\ 0<L<N-K,\ p\in P(K)} \left\{ \frac{N^N}{K^K L^L (N-K-L)^{N-K-L}} \times \right. \tag{4.8}$$

$$\left. \times \frac{(y_{i_1}^2 + \ldots + y_{i_K}^2)^K (y_{j_1}^2 + \ldots + y_{j_L}^2)^L (y_{l_1}^2 + \ldots + y_{l_{N-K-L}}^2)^{N-K-L}}{(y_1^2 + \ldots + y_N^2)^N} \right\}.$$

The main advantage of the test statistic (4.8) is that under the H_0 it does not depend on the unknown value of the parameter σ^2. The null distribution of the LR test statistics $-\ln\lambda_N$ where λ_N is given by (4.8) is derived in the following theorem.

Theorem 1. Let $y_1, \ldots, y_N$ are iid according to the Rayleigh distribution with the unknown scale parameter σ^2. Then the LR test statistics $-\ln\lambda_N$ where λ_N is given by the formula (4.8) has the form

$$-\ln\lambda_N(y) = -\min_{0<K<N-1,\ 0<L<N-K,\ p\in P(K)} \left\{ N\ln N - K\ln K - L\ln L + \right.$$

$$-(N-K-L)\ln(N-K-L) + K\ln\left(\sum_{n=1}^{K} y_{i_n}^2\right) + L\ln\left(\sum_{n=1}^{L} y_{j_n}^2\right) +$$

$$\left. +(N-K-L)\ln\left(\sum_{n=1}^{N-K-L} y_{l_n}^2\right) - N\ln\left(\sum_{n=1}^{N} y_n^2\right)\right\}$$

and it has the same distribution as the random variable

$$V_N = -\min_{0<K<N-1,\ 0<L<N-K,\ p\in P(K)} \left\{ N\ln N - K\ln K - L\ln L + \right.$$

$$-(N-K-L)\ln(N-K-L) + K\ln\left(\sum_{n=1}^{K} u_{i_n}\right) + L\ln\left(\sum_{n=1}^{L} u_{j_n}\right) +$$

$$\left. +(N-K-L)\ln\left(\sum_{n=1}^{N-K-L} u_{l_n}\right) - N\ln\left(\sum_{n=1}^{N} u_n\right)\right\}$$

where $u_1, \ldots, u_N$ are iid according to $Exp(1)$.

Proof. Under H_0, $\frac{x_1^2}{2\sigma_0^2}, \ldots, \frac{x_N^2}{2\sigma_0^2}$ is a random sample from Exp(1). (see (4.5)). The independence of the LR statistics (4.8) on the real value of the scale parameter σ^2 under the null hypothesis completes the proof. □

Remark

The main advantage of the provided distribution of the random variable V_N is the possibility of simulation of the density of the LR statistics $-\ln\lambda_N$ based on the $Exp(1)$ simulations. Power simulations have been conducted in [22].

5 Conclusion

In this paper we illustrated the possibility of divergence testing for reliability engineering. In particular, we have illustrated the importance of the decomposition of divergences, which may provide a form of statistical regularization or optimal statistical procedures. [23] propose a measure of divergence between residual lives of two items that have both survived up to some time t as well as a measure of divergence between past lives. These approaches can bring a potential to new applications of divergences in life time modelling.

Many open problems remain, in particular, generalization and analogical approaches for fuzzy divergences. Some open problems and further research directions related to decompositions are listed in Section 3. We also illustrate how to construct the efficient testing procedure for homogeneity of the scale parameter and the number of components in the Rayleigh mixture.

6 Appendix

6.1 Lambert W function

The Lambert W function is defined to be the multivalued inverse of the complex function $f(y) = ye^y$. As the equation $ye^y = z$ has an infinite number of solutions for each (non-zero) value of $z \in \mathbf{C}$, the Lambert W has an infinite number of branches. Exactly one of these branches is analytic at 0. Usually this branch is referred to as the principal branch of the Lambert W and is denoted by W or W_0. The other branches all have a branch point at 0. These branches are denoted by W_k where $k \in \mathbf{Z} \setminus \{0\}$. The principal branch and the pair of branches W_{-1} and W_1 share an order 2 branch point at $z = -e^{-1}$. A detailed discussion of the branches of the Lambert W can be found in [24]. For more information about the implementation and some computational aspects see [25].

Acknowledgements

Authors are thankful to the anonymous referee for constructive comments which improved paper substantially. Authors also acknowledge the support of Alex Karagrigoriou and Ilia Vonta. The first author acknowledges communication and fruitful discussions with Andrej Pázman. The second author acknowledges the postdoctoral fellowship support.

References

[1] A. Pázman. *Nonlinear statistical models.* Mathematics and its Applications (Dordrecht). 254. Dordrecht: Kluwer Academic Publishers. Bratislava: Ister Science Press Ltd. ix, 1993.

[2] D.W. Coit and T. Jin. Gamma distribution parameter estimation for field reliability data with missing failure times. *IIE Transactions*, 32:1161–1166, 2000.

[3] M. Stehlík. Distributions of exact tests in the exponential family. *Metrika*, 57:145–164, 2003.

[4] F. Rublík. Some tests on exponential populations. *Tatra Mt. Math. Publ.*, 7:229–235, 1996.

[5] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011. ISBN 3-900051-07-0.

[6] M. Stehlík. Exact likelihood ratio scale and homogeneity testing of some loss processes. *Stat. Probab. Lett.*, 76(1):19–26, 2006.

[7] M. Stehlík. Homogeneity and scale testing of generalized gamma distribution. *Reliability Engineering & System Safety*, 93:1809–1813, 2008.

[8] N. Balakrishnan and M. Stehlík. Exact likelihood ratio test of the scale for censored weibull sample. Technical Report 35, IFAS res. report, 2008.

[9] M. Stehlík. Scale testing in small samples with missing time-to-failure information. *International Journal of Reliability, Quality and Safety Engineering*, 16:469–481, 2009.

[10] F. Rublík. On optimality of the lr tests in the sense of exact slopes. *Kybernetika*, 25(1):13–25, 1989.

[11] F. Rublík. On optimality of the lr tests in the sense of exact slopes. ii: Application to individual distributions. *Kybernetika*, 25(2):117–135, 1989.

[12] S.S. Wilks. *Mathematical statistics.* New York and London: John Wiley and Sons, Inc. XIV, 1962.

[13] A. Karagrigoriou and K. Mattheou. On distributional properties and goodness-of-fit tests for generalized measures of divergence. *Commun. Stat., Theory Methods*, 39(3):472–482, 2010.

[14] M Stehlík. Decompositions of information divergences: recent development, open problems and applications. 2012.

[15] F. Liese and I. Vajda. On divergences and informations in statistics and information theory. *IEEE Transactions on Information Theory*, 52(10):4394–4412, 2006.

[16] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.

[17] T. Bengtsson and J. E. Cavanaugh. An improved akaike information criterion for state-space model selection. *Comput. Stat. Data Anal.*, 50(10):2635–2654, 2006.

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

[19] E.L. Lehmann. *Testing statistical hypotheses.* John Wiley & Sons, New York, 1964.

[20] E. Susko. Weighted tests of homogeneity for testing the number of components in a mixture. *Comput. Stat. Data Anal.*, 41(3):367–378, 2002.

[21] M. Stehlík and H. Wagner. Exact likelihood ratio testing for homogeneity of the exponential distribution. *Commun. Stat., Simulation Comput.*, 40(5):663–684, 2011.

[22] M. Stehlík and L. Strelec. On simulation of exact tests in rayleigh and normal families. 2012.

[23] I. Vonta and A. Karagrigoriou. Generalized divergence measures for survival and reliability data. 2012.

[24] R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey, and D.E. Knuth. On the lambert w function. *Adv. Comput. Math.*, 5(4):329–359, 1996.

[25] R.M. Corless, G.H. Gonnet, D.E.G. Hare, and D.J. Jeffrey. Lambert's w function in maple. *THE MAPLE TECHNICAL NEWSLETTER*, 9:12–22, 1993.