National Research University Higher School of Economics

Department of Mathematics
St. Petersburg School of
Physics, Mathematics, and Computer Science

as a manuscript

Ragozin Ilia Andreevich

Goodness-of-fit tests based on some characterizations of Distributions, and their efficiencies

Summary of the PhD thesis
for the purpose of obtaining academic degree
Doctor of Philosophy in Mathematics

Academic superviser
D.Sci., Prof.
Nadezhda Viktorovna Gribkova

Saint Petersburg – 2024

General information

Research Relevance. Statistical hypothesis testing is one of the most important tasks in mathematical statistics. The classical part of hypothesis testing is goodness-of-fit testing, which involves checking whether a sample comes from the particular family of probability distributions. The classical examples are the χ^2-test, the Kolmogorov-Smirnov test, and the Cramer-von Mises-Smirnov test. In last years, the idea of using characterizations of distributions to construct goodness-of-fit tests has become popular. For instance, in the review article [3] a variety of goodness-of-fit criteria for exponential, normal and uniform distributions are described, which are based on characterizations for these families. Such tests are popular because they are convenient to use in practical applications and also, perhaps, due to hidden properties of distributions, expressed precisely in terms of characterizations. The idea of using characterizations in constructing of goodness-of-fit tests emerged in the 1950s and belongs to Yu. V. Linnik [14], but unfortunately until the 1990s there were not tools and the necessary theory for its implementation. Only a few decades later, with the development of the theory of U-statistics, the construction of such criteria became feasible. There are some papers devoted to this direction: Baringhaus and Henze [1, 2], Nikitin and Ponikarov [19], Muliere and Nikitin [10], Litvinova and Nikitin [25, 26], Nikitin and Tchirina [16, 17], [27, 28], Volkova and Nikitin [22], [11], [24] and others.

The significance of construction new goodness-of-fit tests is determined by the fact that one of the fundamental problem in mathematical statistics is the deliberate choice among several available criteria. A characteristic on the basis of which nonparametric tests can be compared and the most appropriate criterion for the given task can be identified is asymptotic efficiency or asymptotic relative efficiency. Of note, the concept of the asymptotic efficiency of a criterion, different variations of which emerged in the late 1940s to early 1950s, is more complicated then the asymptotic efficiency of estimators. Among the three most well-known approaches to calculating asymptotic efficiency, the most convenient and accurate one belongs to Bahadur R.; his approach allows to compare criteria that are not asymptotically normal, making Pitman's approach inapplicable, and, as shown in [15], Bahadur's approach is more sensitive for certain tests than the Hodges-Lehmann approach. Moreover, it is important to note that only single test is sufficient to reject the null hypothesis, and each new criterion that does not contradict it only gradually brings us closer to accepting its validity.

The Aim of the Research. The dissertation research is devoted to the construction and the research of new goodness-of-fit tests based on characterizations, special properties, and the difference of U-empirical Laplace transforms for such properties. The results concern the following: logistic with an arbitrary shift-parameter, exponential, Pareto I type with an arbitrary shape parameter, and Rayleigh distributions. The main purpose of this dissertation is asymptotic comparison of criteria based on the concept of local Bahadur efficiency for close alternatives, including the finding of logarithmic asymptotics of the probability of large deviations under the null hypothesis and the calculation of the asymptotics of the Bahadur exact slopes. Additionally, for the goodness-of-fit criteria for the logistic distribution and the Rayleigh distribution, a problem was posed to compute the empirical power against similar distributions and perform further comparison.

Methodology of the Research. In the dissertation research, methods from the theory of U-statistics, extensively described in [5] and [7], and Bahadur's theory [12, 13], are applied. This theory allows to perform asymptotic comparisons of test statistics [15] based on the fundamental concept of the Bahadur exact slope, the existence and explicit formula of which are proven in Bahadur's fundamental theorem, and local Bahadur efficiency.

The key points are Kullback–Leibler information, logarithmic asymptotics of the probability of large deviations under the null hypothesis, and the law of large numbers under the alternate.

Main Results.

1. Four goodness-of-fit criteria were constructed based on two characterizations of the logistic distribution family with an arbitrary shift parameter. For these criteria, the logarithmic asymptotics of the large deviation function were found, allowing for the calculation of the Bahadur exact slope and local Bahadur efficiency for the considered alternatives. The empirical power was computed for the constructed criteria against distributions that are similar to the logistic distribution. Additionally, the real data from the article [23] was tested.
2. Four goodness-of-fit tests were constructed and studied for the exponential distribution family. Two criteria are based on characterization, and two tests are based on the difference of U-empirical Laplace transforms for a special property from the article [11]. For all criteria, the logarithmic probability of large deviations were found, and the local Bahadur efficiency was calculated.
3. Two goodness-of-fit criteria for the Pareto I type distribution family with an arbitrary shape parameter, based on a certain characterization, were asymptotically researched. For these criteria, the logarithmic asymptotics of the probability of large deviations were computed, and the local Bahadur efficiency was calculated for the considered close alternatives.
4. Five goodness-of-fit tests were constructed for the Rayleigh distribution family, two of which are based on the transformed Desu's characterization, two tests are based on a certain special property, and the remaining one is based on the difference of U-empirical Laplace transforms for this property. Formula for the Kullback-Leibler information was proven for the composite null hypothesis of belonging to the Rayleigh distribution family. An asymptotic comparison based on the values of local Bahadur efficiency was performed for all criteria for the considered close alternatives and the Rice alternative. In the case of integral criteria based on a special property and difference of Laplace transforms, the maximum possible local Bahadur efficiency was found. The empirical power was computed for all criteria against alternative distributions from the article [9].

Novelty of the Research. All the main results of the dissertation research are novel. For the first time, characterizations with random shifts and random stretching/compression are adapted and applied in constructing goodness-of-fit criteria.

Theoretical Significance of the Research. In the dissertation research, logarithmic asymptotics of the probability of large deviations were found for all constructed criteria, and the local Bahadur efficiency was computed for all considered close alternatives. In this research goodness-of-fit criteria for the logistic distribution family with an arbitrary shift parameter and for the Rayleigh distribution family were constucted and asymptotically studied for the first time. Additionally, an asymptotic research of the difference between Rayleigh and Rice distributions were conducted, which is an important problem in statistical radiophysics. The goodness-of-fit criteria for the exponential family and the Pareto I type distribution family outperform previously known criteria for some alternatives and the most efficient in the Bahadur sense. A formula for the Kullback-Leibler information was proved for the composite goodness-of-fit hypothesis for the Rayleigh distribution.

Practical Significance of the Research. All constructed criteria are suitable for practical applications of mathematical statistics, including the testing fit of real data to the particular distribution or the rejection the corresponding goodness-of-fit hypothesis.

Approbation of the Research Results.

The results of the dissertation research were reported at international conferences and seminars:

1. «Stochastic models II» (Saint-Petersburg, 6-8 May, 2019),
2. «The 21st European Young Statisticians Meeting» (Belgrade, Serbia, 29 July-2 August, 2019),
3. «XXXVI International Seminar on Stability Problems for Stochastic Models», (On-line, 22-26 June, 2020),
4. «XXXVI International Seminar on Stability Problems for Stochastic Models» (Petrozavodsk, 21-25 June, 2021),
5. «New trends in mathematical stochastics» (Saint-Petersburg, 30 August–3 September, 2021),
6. «Limit Theorems of Probability Theory and Mathematical Statistics» (Tashkent, 26—28 September 2022).

Publications. The main results of the dissertation research are published in 8 articles: [29–36].

Structure and scope of the thesis. The dissertation research consists of an Introduction, 5 chapters, Conclusion, Bibliography, and Appendices. The total scope of the dissertation is 101 pages. The bibliography contains 96 references (including 8 references to articles published by the author).

Contents

The introduction describes the history of the field, the choice of asymptotic research method, and the general approach to constructing criteria. The structure and contents of the dissertation are presented.

In the first chapter fundamental definitions and formulas are provided, along with necessary auxiliary information from the theory of U-statistics and Bahadur's theory, the theorems of the logarithmic asymptotics of the probability of large deviations are discussed, close alternatives are introduced, and the Kullback-Leibler information is calculated for them.

In the second chapter goodness-of-fit criteria for the logistic distribution family with an arbitrary shift parameter are constructed and studied. These criteria are based on the characterizations by Hu and Lin [8] and the characterization by Ahsanullah-Nevzorov-Yanev [6]. For each characterization, two test statistics of integral type and Kolmogorov type are constructed.

For the first characterization

$$LU_n^L = \int_{-\infty}^{\infty} \left(F_n(t) - U_n^+(t)\right) dF_n(t), \qquad KU_n^L = \sup_t |F_n(t) - U_n^+(t)|,$$

for the second characterization

$$IU_n^L = \int_{-\infty}^{\infty} \left(U_n^+(t) - U_n^-(t)\right) dF_n(t), \qquad QU_n^L = \sup_t |U_n^+(t) - U_n^-(t)|,$$

where $F_n(t)$-usual empirical distribution function, and U_n^+ and U_n^- – the U-empirical distribution functions corresponding to the left and right parts of the second characterization

$$U_n^+(t) = \binom{n}{2}^{-1} \sum_{1\le i<j\le n} \left(1 - e^{\min(0,\min(X_i,X_j)-t)}\right), \qquad U_n^-(t) = \binom{n}{2}^{-1} \sum_{1\le i<j\le n} \left(e^{-\max(0,\max(X_i,X_j)-t)}\right),$$

where $X_1, \ldots, X_n$ – independent identically distributed random variables (i.i.d. r. v.) with an unknown continuous distribution function (d.f.).
At the end of this chapter there is an example of applying the new criteria to real data from the article [23]. Based on the obtained p-values, the hypothesis of belonging to the logistic distribution can not be rejected.

The integral statistics are asymptotically equivalent to the U-statistics of degree 3 with the following centered kernels:

$$\Phi_L(x,y,z) = \frac{1}{2} - \frac{1}{3}\left(g_L(x,y;z) + g_L(y,z;x) + g_L(x,z;y)\right),$$
$$\Phi_I(x,y,z) = \frac{1}{3}\left(g_I(x,y;z) + g_I(y,z;x) + g_I(x,z;y)\right),$$

where

$$g_L(x,y;z) = \left(1 - e^{(\min(x,y)-z)}\right) 1\{\min(x,y) < z\},$$
$$g_I(x,y;z) = e^{-\max(0,\max(x,y)-z)} - (1 - e^{\min(0,\min(x,y)-z)}),\ x,y,z \in \mathbb{R}.$$

The projections of these kernels and the variances of the projections are equal to

$$\Psi_L(t) = -\frac{2}{3}\left(Li_2(-e^t) + t\ln(e^t+1) - \frac{1}{2}\ln^2(e^t+1) + \frac{7e^t+1}{4(e^t+1)}\right),\ \Delta_L^2 = \mathbb{E}\Psi_L^2(X) \approx 0.001899,$$

and

$$\Psi_I(t) = \frac{2}{3}\left(\ln^2(e^t+1) - t\ln(e^t+1) + \frac{\pi^2}{6} - 2\right),\ \Delta_I^2 = \mathbb{E}\Psi_I^2(X) \approx 0.00697.$$

From the calculations carried out in chapter 2, it follows that the kernels Φ_L and Φ_I are non-degenerate, and by Hoeffding's theorem [7] one has

$$\text{As} \quad n\to\infty \quad \sqrt{n}\, LU_n^L \xrightarrow{d} \mathcal{N}\left(0, 9\Delta_L^2\right), \quad \sqrt{n}\, IU_n^L \xrightarrow{d} \mathcal{N}\left(0, 9\Delta_I^2\right).$$

The chapter 2 shows that the kernels are non-degenerate, centered and bounded, therefore by using the results of the study on large deviations of U-statistics presented in [19] the Theorems 1 and 2 were obtained.
Theorem 1. *For every $t > 0$*

$$\lim_{n\to\infty} n^{-1} \ln \mathbb{P}(LU_n^L > t) = h_L(t) \sim -\frac{t^2}{18\Delta_L^2},$$

where h_L is a continuous function for which the asymptotic near zero is essential.

Theorem 2. *For every $t > 0$*

$$\lim_{n\to\infty} n^{-1} \ln \mathbb{P}(IU_n^L > t) = h_I(t),$$

where h_I is a continuous function, such that $h_I(t) \sim -\frac{t^2}{18\Delta_I^2}$ as $t \to 0$.

Kolmogorov-type statistics can be considered as the supremum by t of the family of absolute values for U-statistics with the centered kernels

$$\Phi_K(X,Y;t) = \left(1-e^{(\min(X,Y)-t)}\right)\mathbf{1}\{\min(X,Y) < t\}, \quad \Phi_Q(X,Y;t) = e^{-\max(0,\max(X,Y)-t)} - \left(1-e^{\min(0,\min(X,Y)-t)}\right).$$

In this chapter the projections for each kernel, the variance functions of the projections, and their supremum were computed as

$$\Psi_K(s,t) = \frac{e^{\min(s,t)}(1+e^{-t})}{1+e^{\min(s,t)}} - e^{-t}\ln(e^{\min(s,t)}+1) + \frac{1\{s<t\}(1-e^s(1+2e^{-t}))}{2(1+e^s)} - \frac{e^t}{2(1+e^t)},$$

$$\Delta_K^2(t) := \mathbb{E}_x\Psi_K^2(X,t) = \frac{e^{3t}+8e^{2t}+8e^t-4(e^t+1)(e^t+2)\ln(e^t+1)}{4e^{2t}(e^t+1)^2}, \ \Delta_K^2 = \sup_{t\in\mathbb{R}} \Delta_K^2(t) \overset{t_0=0.3255}{=} 0.02322.$$

For statistics QU_n^L

$$\Psi_Q(s,t) = e^t\left(\ln(1+e^{\max(s,t)}) - \frac{1}{1+e^{\max(s,t)}} - \max(s,t)\right)$$

$$-e^{-t}\left(\frac{(1+e^t)e^{\min(s,t)}}{1+e^{\min(s,t)}} - \ln(1+e^{\min(s,t)})\right) + \frac{e^{3t}+e^{2t}+1\{s<t\}(e^{2t}+e^t+1)(e^s-e^t)}{e^t(1+e^t)(1+e^s)},$$

$$\Delta_Q^2(t) = e^{-t}\left(2te^{3t}-2(t-1)e^{2t}-3e^t+2\right) - 2e^{-2t}\left(e^{4t}-e^{3t}-te^{2t}-e^t+1\right)\ln(1+e^t) - 2\ln^2(1+e^t),$$

and its supremum is attained at $t_0 = 0$ and equals $1 - 2\ln^2(2) \approx 0.00393$.

Therefore, the kernels are non-degenerate. Since the families of kernels are centered and bounded, applying the theorem from [21] on the logarithmic asymptotics of the probability of the large deviation under the null hypothesis H_0 for the U-empirical Kolmogorov-type statistics to the following relation we obtained the following results in this chapter.

Theorem 3. *For every $z > 0$*

$$\lim_{n\to\infty} n^{-1}\ln\mathbb{P}\left\{KU_n^L > z\right\} = w_K(z) \sim -\frac{z^2}{8\Delta_K^2},$$

where w_K is a continuous function, for which the asymptotic behavior towards zero is essential.

Theorem 4. *For every $z > 0$*

$$\lim_{n\to\infty} n^{-1}\ln\mathbb{P}\left\{QU_n^L > z\right\} = h_Q(z) \sim -\frac{z^2}{8\Delta_Q^2},$$

where h_Q is a continuous function for which the asymptotic behavior towards zero is essential.

Then, for each of the statistics, the Bahadur exact slope $c(\theta)$ is computed [12,13], and the doubled Kullback-Leibler information, which is an upper bound for $c(\theta)$, is calculated for the following alternatives.

1. Scale alternative with the density

$$f_1(x,\theta) = \frac{e^{\theta + xe^{\theta}}}{(1+e^{xe^{\theta}})^2};$$

2. Generalized hyperbolic cosine alternative with the density

$$f_2(x,\theta) = \frac{\Gamma(\theta+2)}{\Gamma^2(\frac{\theta}{2}+1)} \frac{e^{\left(x+\frac{\theta x}{2}\right)}}{(1+e^x)^{\theta+2}};$$

3. Sine-alternative with the density for small θ

$$f_3(x,\theta) = l(x) - 2\pi\theta\cos(2\pi L(x))l(x),$$

where $L(x) = (1+\exp(-x))^{-1}$, $l(x) = \frac{e^{-x}}{(1+e^{-x})^2}$ $x \in \mathbb{R}$.

Table 1. The Kullback-Leibler information for alternatives $f_i(x,\theta)$, $i{=}\overline{1,3}$ as $\theta \to 0$

	Altertatives		
	f_1	f_2	f_3
$2K(\theta)$	$1.4300 \cdot \theta^2$	$0.1775 \cdot \theta^2$	$19.7392 \cdot \theta^2$

In this chapter, the local Bahadur efficiency is calculated using the formula, which values are collected in the table 2 below. The next step included the empirical power calculation for the standard normal

Table 2. Comparative of local Bahadur efficiencies for the test statistics

	Test statistics							
	LU_n^L		IU_n^L		KU_n^L		QU_n^L	
Alternatives	$c(\theta)$	eff	$c(\theta)$	eff	$c(\theta)$	eff	$c(\theta)$	eff
f_1	$1.229 \cdot \theta^2$	0.860	$1.339 \cdot \theta^2$	**0.937**	$0.505 \cdot \theta^2$	0.353	$0.954 \cdot \theta^2$	0.667
f_2	$0.141 \cdot \theta^2$	0.791	$0.153 \cdot \theta^2$	**0.864**	$0.051 \cdot \theta^2$	0.288	$0.101 \cdot \theta^2$	0.566
f_3	$15.384 \cdot \theta^2$	0.779	$16.76 \cdot \theta^2$	0.849	$15.791 \cdot \theta^2$	0.800	$19.253 \cdot \theta^2$	**0.975**

distribution and the logistic distribution $Logist(0, e^{\theta})$ for sample sizes n=20 and n=50 at 0.05 and 0.01 significance levels. The results are presented in table 3.

At the end of the chapter 2 IU_n^L and QU_n^L were applied to the real data from the article [23] and generated data from the standard Cauchy distribution. Based on the p-values in table 4, it is seen that the data from the first set confidently support the hypothesis H_0, and this corresponds to the conclusion from the paper [23], and our tests surely reject the null hypothesis for the second data set as the p-values are convincing enough.

Table 3. Power for the statistics IU_n^L, QU_n^L, LU_n^L and KU_n^L

			IU_n^L		QU_n^L		LU_n^L		KU_n^L	
$Logist(0, e^{\theta})$	θ	n	5%	1%	5%	1%	5%	1%	5%	1%
	$\theta = 0.75$	n=20	**0.963**	**0.839**	0.953	0.806	0.911	0.696	0.883	0.668
		n=50	1	1	1	0.999	1	0.999	1	0.993
	$\theta = 0.5$	n=20	0.621	0.286	**0.673**	**0.400**	0.506	0.264	0.542	0.277
		n=50	**0.980**	**0.917**	0.966	0.892	0.962	0.84	0.911	0.706
	$\theta = 0.25$	n=20	0.203	**0.09**	**0.24**	0.089	0.156	0.055	0.202	0.067
		n=50	**0.484**	**0.252**	0.463	0.244	0.423	0.188	0.372	0.149
Norm(0,1)		20	**0.758**	0.415	0.602	0.319	0.749	**0.461**	0.589	0.294
		50	**0.997**	**0.974**	0.995	0.935	0.992	0.949	0.960	0.813

Table 4. Values of statistics IU_n^L, QU_n^L for the data sets

	The first data set		The second data set	
Test	Test statistic value	p-value	Test statistic value	p-value
IU	0.002	p=0.864	-0.028	p=0.002
QU	0.119	p=0.209	0.146	< 0.002

In the third chapter four goodness-of-fit tests are considered for the exponential family of distributions. Two of them are based on the characterizations by Ahsanullah and Anis from the work [4], the other two are constructed on the idea of considering the difference of U-empirical Laplace transforms for a special property from the [11]. Special property can be formulated as follows, that the distribution of the quotient of two independent exponential random variables is Fisher distribution (2,2). In the case of characterizations, statistics of integral type and Kolmogorov-type (supremum) are constructed as follows:

$$IU_n^E = \int_0^{\infty} (LU_n(t) - RU_n(t))\, dF_n(t), \quad KU_n^E = \sup_t |LU_n(t) - RU_n(t)|,$$

where $F_n(t)$ – usual empirical d.f., and $LU_n(t)$ and $RU_n(t)$ are the U-empirical d.f. corresponding to the left and right parts of the characterization

$$LU_n(t) = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} \mathbb{1}\left\{\max(X_i, X_j) < t\right\}, \quad RU_n(t) = \binom{n}{3}^{-1} \sum_{1 \le i < j < k \le n} \mathbb{1}\left\{\min(X_i, X_j) + X_k < t\right\},$$

where $X_1, \ldots, X_n$ – i.i.d.r.v's. with some d.f.

The integral statistic with exponential weight and the Kolmogorov-type statistic are considered as criteria based on the difference U-empirical Laplace transforms

$$L_n^E(a) = \int_0^{\infty} (LL(t) - RL(t))\, e^{-a \cdot t} dt, \qquad KL_n^E = \sup_{t>0} |LL(t) - RL(t)|,$$

where

$$LL(t) = \binom{n}{2}^{-1} \sum_{1\le i<j\le n} e^{-t\cdot\frac{X_i}{X_j}}, \qquad RL(t) = \int\limits_0^\infty \frac{1}{(1+x)^2}\cdot e^{-t\cdot x}dx = 1 + t\,e^t\,Ei(-t).$$

The integral statistic IU_n^E is asymptotically equivalent to the U-statistic of degree 4 with the centered kernel $\Phi_I(X,Y,Z,W)$:

$$\Phi_I(X_1,X_2,X_3,X_4) = \frac{1}{12}\left(\sum_{\substack{i,j,k\in(1,2,3,4)\\ i\ne j\ne k}} 1\{\max(X_i,X_j)<X_k\}\right) - \frac{1}{12}\Big(1\{\min(X_1,X_2)+X_3<X_4\}$$
$$+1\{\min(X_1,X_2)+X_4<X_3\}+1\{\min(X_2,X_3)+X_4<X_1\}+1\{\min(X_2,X_3)+X_1<X_4\}+$$
$$+1\{\min(X_3,X_4)+X_1<X_2\}+1\{\min(X_3,X_4)+X_2<X_1\}+1\{\min(X_1,X_3)+X_2<X_4\}+$$
$$+1\{\min(X_1,X_3)+X_4<X_2\}+1\{\min(X_1,X_4)+X_2<X_3\}+1\{\min(X_1,X_4)+X_3<X_2\}+$$
$$+1\{\min(X_2,X_4)+X_1<X_3\}+1\{\min(X_2,X_4)+X_3<X_2\}\Big)$$

with the following projection

$$\Psi_I(t) = -\frac{3e^{-2t}}{8}+\frac{e^{-t}}{3}-\frac{1}{24}.$$

The variance of this projection is

$$\Delta_I^2 = \mathbb{E}\{\Psi_I^2(X)\} = \frac{1}{1080}.$$

Consequently, the kernels Φ_I is non-degenerate, hence by Hoeffding's theorem [7] as $n\to\infty$

$$\sqrt{n}IU_n^E \xrightarrow{d} \mathcal{N}\left(0,16\Delta_I^2\right).$$

Since the kernel Φ_I is non-degenerate, centered and bounded, therefore we can use the result of the large deviations for U-statistics from [19]. As a result, in chapter 3 the following theorem is formulated.

Theorem 5. *For any $t>0$*

$$\lim_{n\to\infty} n^{-1}\ln\left(\mathbb{P}(IU_n^E>t)\right) = h(t) \sim -\frac{t^2}{32\Delta_I^2},$$

where h is a continuous function for which the asymptotic near zero is essential.

Kolmogorov-type statistics KU_n^E can be considered as the supremum by t of the family of absolute values for U-statistics with the centered kernels

$$\begin{aligned}\Phi_K(X,Y,Z;t) =& \frac{1}{3}\left(1\{\max(X,Y)<t\}+1\{\max(X,Z)<t\}+1\{\max(Y,Z)<t\}\right)\\ &-\frac{1}{3}\left(1\{\min(X,Y)+Z<t\}+1\{\min(X,Z)+Y<t\}+1\{\min(Y,Z)+X<t\}\right)\end{aligned}$$

and the projections of this kernel is

$$\Psi_K(s,t) = \frac{1}{3}\left((1-e^{-t})^2 - 1\{s<t\}\left(2e^{-s}-1-e^{-2(t-s)}\right)\right)+\frac{2}{3}\left(\min(s,t)e^{-t}-1+e^{-\min(s,t)}\right).$$

The variance function of the projection and its supremum are equal to

$$\Delta_K^2(t) = \mathbb{E}\left\{\Psi_K^2(X,t)\right\} = \frac{4}{27}e^{-4t}(e^t-1)^3, \quad \Delta_K^2 = \sum_{t>0}\Delta_K^2(t) \overset{t_0=\ln(4)}{=} \frac{1}{64}.$$

Therefore, as the kernels are non-degenerate, and the families of kernels are centered and bounded, applying the theorem from [21] on the logarithmic asymptotics of the probabilities of the large deviation under the null hypothesis H_0 for the U-empirical Kolmogorov-type statistics to the following relation the next result is obtained.

Theorem 6. *For every $z > 0$*

$$\lim_{n\to\infty} n^{-1}\ln \mathbb{P}\left\{KU_n^E > z\right\} = w(z) \sim -\frac{z^2}{8\Delta_K^2},$$

where w is a continuous function for which the asymptotic behavior towards zero is essential.

The integral statistic $L_n^E(a)$ is asymptotically equivalent to the U-statistic of degree 2 with the centered kernels

$$\Phi_L(x,y;a) = \frac{1}{2}\left(\frac{1}{a+\frac{x}{y}} + \frac{1}{a+\frac{y}{x}}\right) - \frac{a-1-\ln(a)}{(a-1)^2}.$$

The projection of the kernel is found to be

$$\Psi_L(s;a) = \mathbb{E}\left(\Phi_L(X,Y;a)|Y=s\right) = \frac{1}{2}\left(\frac{a+s\,e^{\frac{s}{a}}Ei(-\frac{s}{a})}{a^2} - s\,e^{as}Ei(-as)\right) - \frac{a-1-\ln(a)}{(a-1)^2},$$

The variance function of the projection, and some value of the variance can be calculated as

$$\Delta_L^2(a) = \mathbb{E}(\Psi_L^2(X;a)), \Delta_L^2(2) = 0.00014,\ \Delta_L^2(0.5) = 0.00228,\ \Delta_L^2(3) = 0.00014,\ \Delta_L^2(4) = 0.00011,$$

hence the kernel Φ_L is non-degenerate. After that, we applied the Hoeffding's theorem [7]

$$\text{as} \quad n\to\infty: \qquad \sqrt{n}\,L_n^E(a) \xrightarrow{d} N(0, 4\Delta_L^2(a)).$$

Since the kernel Φ_L is nondegenerate, centered, and bounded, we used the results of the large deviations of U-statistics from [19] to obtain the following.

Theorem 7. *For any $t > 0$*

$$\lim_{n\to\infty} n^{-1}\ln\left(\mathbb{P}(LE_n > t)\right) = h(t) \sim -\frac{t^2}{4\Delta_L^2},$$

where h is a continuous function for which the asymptotic behavior towards zero is essential.

Kolmogorov-type statistic KL_n^E can be considered as the supremum by t of the family of absolute values for U-statistics with the centered kernels

$$\Phi_{KL}(X,Y;t) = \frac{1}{2}\left(e^{-t\frac{X}{Y}} + e^{-t\frac{Y}{X}}\right) - (1+te^tEi(-t)).$$

In this chapter we found the projection of this kernel

$$\Psi_{KL}(s;t) = \mathbb{E}\left(\Phi_{KL}(X,Y;t)|Y=s\right) = \frac{1}{2}\left(\frac{s}{s+t} + 2\sqrt{st}\,K_1(2\sqrt{st})\right) - (1+te^tEi(-t)),$$

where $K_1(\cdot)$ is the modified Bessel function of the second kind.
The variance function of the projection cannot be expressed explicitly, however, its supremum can be found as $\Delta^2_{KL}(t) = \mathbb{E}\left(\Psi_{KL}(X;t)\right)$, $\quad \Delta^2_{KL} = \sup_{t>0}\Delta^2_{KL}(t) \overset{t_0=0.551}{=} 0.00507$. Hence kernel Φ_{KL} is non-degenerate. Since the family of kernels is centered and bounded, we can apply the theorem from [21] on the logarithmic asymptotics of the probabilities of the large deviation under the null hypothesis H_0 for the U-empirical Kolmogorov-type statistics to the following relation to prove the following.
Theorem 8. *For any* $z > 0$

$$\lim_{n\to\infty} n^{-1}\ln\mathbb{P}\left\{KL_n^E > z\right\} = g(z) \sim -\frac{z^2}{8\Delta^2_{KL}},$$

where g is a continuous function for which the asymptotic near zero is essential.
In this chapter the following close alternatives $f_i(x,\theta)$, $i = 1, ..., 6$, $x > 0$ are considered and the Kullback-Leibler information for them is calculated.

1. Weibull alternative with density

$$f_1(x,\theta) = (1+\theta)x^\theta e^{-x^{1+\theta}},\ \theta \geq 0,\ x \geq 0;$$

2. Lehmann alternative with density

$$f_2(x,\theta) = (1+\theta(1-e^{-x}))e^{(-x-\theta(e^{-x}-1+x))},\ \theta \geq 0,\ x \geq 0;$$

3. Linear failure rate alternative

$$f_3(x,\theta) = (1+\theta x)e^{-x-\frac{\theta x^2}{2}},\ \theta \geq 0,\ x \geq 0;$$

4. Gamma alternative with density

$$f_4(x,\theta) = \frac{x^\theta e^{-x}}{\Gamma(\theta+1)},\ \theta \geq 0,\ x \geq 0;$$

5. Verhulst alternative with density

$$f_5(x,\theta) = (1+\theta)e^{-x}(1-e^{-x})^\theta,\ \theta \geq 0,\ x \geq 0;$$

6. Exponential mixture with negative weight alternative with density

$$f_6(x,\theta) = (1+\theta)e^{-x} - \theta\beta e^{-\beta x},\ \theta \in [0, \frac{1}{\beta-1}],\ \beta \geq 1,\ x \geq 0.$$

The following formula for the Kullback-Leibler information holds for such alternatives [16]:

$$2K(\theta) \sim \left(\int_0^\infty e^x f'_\theta(x,0)^2dx - \left(\int_0^\infty x f'_\theta(x,0)dx\right)^2\right)\cdot\theta^2,\ \theta \to 0.$$

Table 5. The Kullback-Leibler information as $\theta \to 0$

	Alternatives					
	f_1	f_2	f_3	f_4	f_5	f_6
$2K(\theta)$	$\frac{\pi^2\theta^2}{6}$	$\frac{\theta^2}{12}$	θ^2	$\left(\frac{\pi^2}{6}-1\right)\theta^2$	$\left(\frac{\pi^2}{3}-\frac{\pi^4}{36}\right)\theta^2$	$\frac{(\beta-1)^4}{\beta^2(2\beta-1)}\cdot\theta^2$

The obtained asymptotics are collected in the table 5.

Using Theorems 4–8 and improved formulas for Bahadur exact slopes from the articles [11] and [20], we calculated the local Bahadur efficiency for our statistics applying the formulas presented below.
For the integral statistics

$$eff_{IU,j} = \lim_{\theta\to 0}\frac{\left(\int\limits_0^\infty \Psi_I(x) f'_{\theta,j}(x,0)dx\right)^2}{\Delta_I^2\cdot 2K_j(\theta)}, \quad eff_{L,j}(a) = \lim_{\theta\to 0}\frac{\left(\int\limits_0^\infty \Psi_L(x;a) f'_{\theta,j}(x,0)dx\right)^2}{\Delta_L^2(a)\cdot 2K_j(\theta)},$$

and for the Kolmogorov-type statistics

$$eff_{KU,j} = \lim_{\theta\to 0}\frac{\sup\limits_{t\geq 0}\left(\int\limits_0^\infty \Psi_{KU}(x;t) f'_{\theta,j}(x,0)dx\right)^2}{\Delta_{KU}^2\cdot 2K_j(\theta)}, \quad eff_{KL,j} = \lim_{\theta\to 0}\frac{\sup\limits_{t\geq 0}\left(\int\limits_0^\infty \Psi_{KL}(x;t) f'_{\theta,j}(x,0)dx\right)^2}{\Delta_I^2\cdot 2K_j(\theta)}.$$

All calculated values of local Bahadur efficiency are given in table 6.

Table 6. Local Bahadur efficiency

Alternatives	IU_n^E	KU_n^E	$\max(L_n^E)$	KL_n^E
f_1	0.750	0.277	0.812	0.795
f_2	**0.625**	0.342	0.56	0.53
f_3	0.208	0.155	**0.214**	0.202
f_4	0.796	0.267	**0.923**	0.92
f_5	0.803	0.268	**0.941**	0.938
f_6, b=3	**0.844**	0.396	0.761	0.727
f_6, b=4	**0.933**	0.397	0.875	0.842
f_6, b=5	**0.957**	0.377	0.938	0.911
f_6, b=6	0.947	0.355	**0.972**	0.952
f_6, b=7	0.922	0.331	**0.986**	0.974
f_6, b=8	0.889	0.307	**0.989**	0.985

The fourth chapter includes information about two goodness-of-fit tests for the Pareto I type distribution family with arbitrary shape parameter, based on characterization derived from Ahsanullah-Anis characterization for the exponential distribution and using the idea of random compressions and stretches. This characterization can be formulated as follows.
Let Y_1, Y_2, Y – i.i.d.r.v's. with continuous distribution function G, such that $G(1) = 0$, $G(x) > 0$ for all

x>1, infinitely differentiable and g(x) – the corresponding density. $max(Y_1,Y_2)$ and $min(Y_1,Y_2)\cdot Y$ are equally distributed if and only if

$$G(x)=1-x^{-\lambda},\quad x\geq 1,\quad \lambda>0.$$

In chapter 4 the integral statistic and the Kolmogorov-type statistic are considered

$$IU_n^P=\int\limits_1^\infty (LU_n(t)-RU_n(t))\,dG_n(t),\qquad KU_n^P=\sup_{t\geq 1}|LU_n(t)-RU_n(t)|,$$

where $Y_1,\dots,Y_n$ – i.i.d.r.v's with d.f. G, $G_n(t)$ is an usual empirical d.f., namely

$$G_n(t)=n^{-1}\sum_{i=1}^n 1\{Y_i<t\},$$

and $LU_n(t)$ and $RU_n(t)$ are the U-empirical d.f. corresponding to the left and the right parts of the characterization

$$LU_n(t)=\binom{n}{2}^{-1}\sum_{1\leq i<j\leq n}1\left\{\max(Y_i,Y_j)<t\right\},\qquad RU_n(t)=\binom{n}{3}^{-1}\sum_{1\leq i<j<k\leq n}1\left\{\min(Y_i,Y_j)\cdot Y_k<t\right\}.$$

In the dissertation research it is shown that integral statistic IU_n^P is asymptotically equivalent to the U-statistic of degree 4 with the centered kernels

$$\Phi_I(Y_1,Y_2,Y_3,Y_4)=\frac{1}{12}\left(\sum_{\substack{i,j,k\in(1,2,3,4)\\ i\neq j\neq k}}1\{\max(Y_i,Y_j)<Y_k\}\right)$$

$$-\frac{1}{12}(1\{\min(Y_1,Y_2)\cdot Y_3<Y_4\}+1\{\min(Y_1,Y_2)\cdot Y_4<Y_3\}$$
$$+1\{\min(Y_2,Y_3)\cdot Y_4<Y_1\}+1\{\min(Y_2,Y_3)\cdot Y_1<Y_4\}$$
$$+1\{\min(Y_3,Y_4)\cdot Y_1<Y_2\}+1\{\min(Y_3,Y_4)\cdot Y_2<Y_1\}$$
$$+1\{\min(Y_1,Y_3)\cdot Y_2<Y_4\}+1\{\min(Y_1,Y_3)\cdot Y_4<Y_2\}$$
$$+1\{\min(Y_1,Y_4)\cdot Y_2<Y_3\}+1\{\min(Y_1,Y_4)\cdot Y_3<Y_2\}$$
$$+1\{\min(Y_2,Y_4)\cdot Y_1<Y_3\}+1\{\min(Y_2,Y_4)\cdot Y_3<Y_2\}),$$

with the projection equals to

$$\Psi_I(t)=-\frac{3}{8t^2}+\frac{1}{3t}-\frac{1}{24}.$$

The variance of the projection was calculated in chapter 4 and equals to

$$\Delta_I^2=\mathbb{E}\{\Psi_I^2(X)\}=\frac{1}{1080}.$$

Therefore, statistic IU_n^P is non-degenerate, and then, according to the Hoeffding theorem [7] as $n\to\infty$

$$\sqrt{n}IU_n^P\xrightarrow{d}\mathcal{N}\left(0,16\Delta_I^2\right).$$

Since the kernel Φ_I is non-degenerate, centered and bounded, we used the results of large deviations of U-statistics from [19] to obtain the following result.

Theorem 9. *For any $t > 0$*

$$\lim_{n\to\infty} n^{-1} \ln \left(\mathbb{P}(IU_n^P > t)\right) = h(t),$$

where h is a continuous function, such that

$$h(t) \sim -\frac{t^2}{32\Delta_I^2}, \quad \text{as } t \to 0.$$

Kolmogorov-type statistic KU_n^P can be considered as the supremum by t of the family of absolute values for U-statistics with the centered kernels and the following projections:

$$\Psi_K(s,t) = \frac{1}{3}\left(\left(1-\frac{1}{t}\right)^2 + 1\{s<t\}\left(1-\frac{2}{s}+\frac{s^2}{t^2}\right)\right)$$
$$+\frac{2}{3}\left(\frac{\ln(\min(s,t))}{t} + \frac{1}{\min(s,t)} - 1\right).$$

In this chapter the variance of this projection and its supremum were found to be

$$\Delta_K^2(t) = \frac{4(t-1)^3}{27t^4}, \quad \Delta_K^2 = \sup_{t\geq 1} \Delta_K^2(t) \overset{t_0=4}{=} \frac{1}{64}.$$

Consequently the family of kernels is non-degenerate, and also centered, and bounded by the considered characterization, we applied the results from [21] on the logarithmic asymptotics of the probabilities of the large deviation under the null hypothesis H_0 for the U-empirical Kolmogorov-type statistics to the following relation to obtain the next result.

Theorem 10. *For any $z > 0$*

$$\lim_{n\to\infty} n^{-1} \ln \mathbb{P}\left\{KU_n^P > z\right\} = k(z),$$

where k is a continuous function, such that

$$k(z) \sim -\frac{z^2}{18\Delta_K^2}, \quad \text{as } z \to 0.$$

In this chapter we described the alternatives $f_i(x,\theta)$, $x \geq 1$, $i = 1,\ldots,6$.

1. Shift alternative with the density

$$F_1(x,\theta) = 1 - \frac{\theta+1}{x+\theta}, \quad f_1(x,\theta) = \frac{1+\theta}{(x+\theta)^2}, \quad \theta \geq 0, \;\; x \geq 1;$$

2. Log-Weibull alternative with the probability function

$$F_2(x,\theta) = 1 - e^{-(\ln(x))^{\theta+1}}, \quad x \geq 1, \;\; \theta \in (0,1);$$

3. Lehmann alternative with the probability function

$$F_3(x,\theta) = F^{1+\theta}(x) = (1-x^{-1})^{1+\theta}, \quad x \geq 1, \;\; \theta > 0;$$

4. Pareto IV type alternative with the probability function

$$F_4(x,\theta) = P_{IV}\left(1, 1, \frac{1}{1+\theta}, 1\right)(x) = 1 - \frac{1}{1+(x-1)^{1+\theta}}, \quad x \geq 1, \ \theta \geq 0;$$

5. Sine alternative with the probability function

$$F_5(x,\theta) = F(x) - \theta \sin(\pi F(X)), \quad \theta \in [0, \frac{1}{\pi}], \ x \geq 1;$$

6. Mixture alternative of two Pareto distributions with the probability function

$$F_6(x,\theta) = (1-\theta)\frac{x-1}{x} + \theta(1 - x^{-\beta}), \quad x \geq 1, \ \beta > 1, \ \theta \in (0,1).$$

Since the alternatives $f_i(x,\theta)$ $i = 1, ..., 6$ satisfy the regularity conditions (see [12], §4), the following expression for the Kullback-Leibler information was obtained [18]:

$$2K(\theta) \sim \left\{ \int_1^\infty x^2 f'_\theta(x,0)^2 dx - \Big[\int_1^\infty \ln(x) f'_\theta(x,0) dx \Big]^2 \right\} \cdot \theta^2.$$

We collected the values of the Kullback-Leibler information in the table 7 below.

Table 7. The Kullback-Leibler information as $\theta \to 0$

	Alternatives					
	f_1	f_2	f_3	f_4	f_5	f_6
$2K(\theta)$	$\frac{\theta^2}{12}$	$\frac{\pi^2}{6}\theta^2$	$\left(\frac{\pi^2}{3} - \frac{\pi^4}{36}\right)\theta^2$	$\frac{1}{36}(3+4\pi^2)\cdot\theta^2$	$\left(\frac{\pi^2}{2} - Si^2(\pi)\right)\cdot\theta^2$	$\frac{(\beta-1)^4}{\beta^2(2\beta-1)}\cdot\theta^2$

The local Bahadur efficiency for the considered alternatives $f_i(x,\theta)$ $i = 1, ..., 6$ for test statistics using the formulas from [20], [11] were calculated as

$$eff_{IU,j} = \lim_{\theta\to 0} \frac{\left(\int_0^\infty \Psi_I(x) f'_{\theta,j}(x,0)dx\right)^2}{\Delta_I^2 \cdot 2K_j(\theta)}, \quad eff_{KU,j} = \lim_{\theta\to 0} \frac{\sup_{t\geq 1}\left(\int_0^\infty \Psi_{KU}(x;t) f'_{\theta,j}(x,0)dx\right)^2}{\Delta_{KU}^2 \cdot 2K_j(\theta)}, \ j = 1, ..., 6$$

and the values are collected in the table 8 below.

Table 8. Local Bahadur efficiency for IU_n^P and KU_n^P

	f_1	f_2	f_3	f_4	f_5	f_6, $b=3$	f_6, $b=4$	f_6, $b=5$	f_6, $b=6$
IU_n^P	0.625	0.750	0.804	0.894	0.505	0.844	0.933	0.957	0.947
KU_n^P	0.342	0.277	0.268	0.311	0.394	0.396	0.397	0.377	0.355

The fifth chapter gives information about five goodness-of-fit tests for the Rayleigh distribution family, two of them are based on the transformed Desu characterization, two criteria are based on a special property, and the last one is based on the difference of Laplace U-empirical transformations for this property. We formulate the characterization.

Let X and Y – i.i.d. positive r.v's with a continuous d.f. R. Then

$$X \stackrel{d}{=} \sqrt{2}\,\min(X,Y),$$

if and only if R belongs to the Rayleigh distribution family with arbitrary scale parameter $\sigma > 0$ with density $r(x,\sigma)$, $\sigma > 0$, $x \geq 0$, where

$$r(x,\sigma) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}.$$

Now we formulate the special property.

Let X, Y – i.i.d. r.v's with d. f. Rayleigh(σ), then its quotient has a distribution determined by the following probability distribution function and density:

$$q(x) = \frac{2x}{(1+x^2)^2}, \quad Q(x) = \frac{x^2}{1+x^2}, \quad x \geq 0.$$

We computed the Laplace transform for density in the right part

$$L_2(t) = \int_0^\infty e^{-t\cdot x} q(x)dx = 1 - \frac{\pi \cdot t \cdot \cos(t)}{2} - tCi(t)\cdot\sin(t) + t\cos(t)\cdot Si(t),$$

where $Ci(\cdot), Si(\cdot)$ are the integral Cosine and Sine.

Let $X_1, ..., X_n$ are i. i. d. observations with d. f. R. We constructed the integral statistic and the Kolmogorov-type statistic based on characterization

$$IU_{1,n}^R = \int_0^\infty \left(F_n(t) - U_{1,n}(t)\right) dF_n(t), \qquad KU_{1,n}^R = \sup_{t\geq 0} |F_n(t) - U_{2,n}(t)|,$$

where $F_n(t)$ is a usual empirical d. f., and $U_{1,n}(t)$- U-empirical d. f. corresponding to the right part of the characterization:

$$U_{1,n}(t) = \binom{n}{2}^{-1} \sum_{1\leq i<j\leq n} 1\left\{\sqrt{2}\,\min(X_i, X_j) < t\right\}.$$

We introduce the integral weighted statistic and the Kolmogorov-type statistic, based on the special property

$$IU_{2,n,\sigma}^R = \int_0^\infty \left(U_{2,n}(t) - \frac{t^2}{1+t^2}\right) \sigma^2\, t\, e^{-\frac{\sigma^2 t^2}{2}}\, dt, \qquad KU_{2,n}^R = \sup_{t\geq 0}\left|U_{2,n}(t) - \frac{t^2}{1+t^2}\right|,$$

where

$$U_{2,n}(t) = \binom{n}{2}^{-1} \left(\sum_{1\leq i<j\leq n} \left(\frac{1\{\frac{X_i}{X_j} < t\} + 1\{\frac{X_j}{X_i} < t\}}{2}\right)\right).$$

And we introduce the integral weighted statistic based on the difference of Laplace transforms for this property

$$L_n(a) = \int_0^\infty \left(L_1(t) - L_2(t)\right) e^{-a\cdot t} dt,$$

where $L_1(t)$- U-empirical Laplace transform for $\frac{X}{Y}$

$$L_1(t) = (c_n^2)^{-1} \sum_{1 \le i < l \le n} e^{-t \cdot \frac{X_i}{X_j}}$$

and $L_2(t)$ is obtained previously.

The integral statistic $IU_{1,n}$ is asymptotically equivalent to the U-statistics of degree 3 with the following centered kernel:

$$\Phi_1(x, y, z) = \frac{1}{3}\left(g(x, y; z) + g(y, z; x) + g(x, z; y)\right),$$

where $g(x, y; z) = \frac{1}{2} - 1\left\{\sqrt{2}\,\min(x, y) < z\right\}$. The following projection of this kernel is as follows:

$$\Psi_1(t) = -\frac{1}{18} + \frac{1}{3}e^{-\frac{t^2}{2}} - \frac{4}{9}e^{-\frac{3t^2}{2}}, \quad t \ge 0.$$

The variance of the projection can be presented as

$$\Delta_1^2 = \mathbf{E}\Psi_1^2(X) \approx 0.00291,$$

hence the kernel Φ_1 is non-degenerate, and by the Hoeffding's theorem [7]

$$\sqrt{n}\, IU_{1,n} \xrightarrow{d} \mathcal{N}\left(0, 9\Delta_1^2\right), \quad n \to \infty.$$

The kernel are non-degenerate, centered and bounded, therefore we used the results of large deviations of U-statistics presented in [19] to obtain the following.

Theorem 11. *For any* $t > 0$

$$\lim_{n \to \infty} n^{-1} \ln \mathbf{P}(IU_{1,n} > t) = h_1(t),$$

where h *is a continuous function, such that* $h_1(t) \sim -\frac{t^2}{18\,\Delta_1^2}$, *as* $t \to 0$.

The integral statistic $IU_{2,n}^R$ is equivalent to the U-statistic of degree 2 with centered kernel

$$\Phi_I(x, y; \sigma) = \frac{e^{-\frac{\sigma^2 x^2}{2y^2}} + e^{-\frac{\sigma^2 y^2}{2x^2}}}{2} - \left(1 + \frac{1}{2}\, e^{\frac{\sigma^2}{2}}\, \sigma^2\, Ei\left(-\frac{\sigma^2}{2}\right)\right)$$

$$= \frac{e^{-\frac{\sigma^2 x^2}{2y^2}} + e^{-\frac{\sigma^2 y^2}{2x^2}}}{2} - c(\sigma),$$

where $Ei(\cdot)$ is the exponential integral .

The projection of this kernel is

$$\Psi_I(t; \sigma) = \frac{1}{2}\left(\frac{t^2}{\sigma^2 + t^2} + \sigma\, t\, K_1(\sigma \cdot t)\right) - c(\sigma), \quad t \ge 0,$$

where $K_1(t)$ is the modified Bessel function of the second kind.

The variance function of the projection cannot be expressed explicitly but can be presented as follows:

$$\Delta_I^2(\sigma) = \mathbf{E}\Psi_I^2(X; \sigma) = \int_0^\infty \Psi_I^2(x; \sigma)\, x\, e^{-\frac{x^2}{2}}\, dx > 0,$$

however, for every $\sigma > 0$ we can calculate the value of the variance by the numeral methods.
For example, the variance of the projection for $\sigma = 2$ is $IU_{2,n} := IU_{2,n}(1)$: $\Delta_I^2 = \Delta_I^2(1) \approx 0.000314$, hence Φ_I is non-degenerate, and then by the Hoeffding's theorem [7] as $n \to \infty$

$$\sqrt{n}\, IU^R_{n,\sigma^2} \xrightarrow{d} \mathcal{N}\left(0, 4\Delta_I^2(\sigma)\right).$$

Since the kernel Φ_I is non-degenerate, centered, and bounded, we applied the results of the large deviations of U-statistics from [19] to obtain the following.

Theorem 12. *For any $t > 0$*

$$\lim_{n\to\infty} n^{-1} \ln\left(\mathbf{P}(IU^R_{n,\sigma^2} > t)\right) = h_I(t,\sigma) \sim -\frac{t^2}{8\Delta_I^2(\sigma)} = -h_I(t,\sigma),$$

where h_I is a continuous function for which the asymptotic near zero is essential.

Kolmogorov-type statistics $KU_{1,n}$ and $KU_{2,n}$ can be considered as the supremum by t of the family of absolute values for U-statistics with the centered kernels

$$\Phi_1(x,y;t) = \frac{1}{2}\left(1\{x < t\} + 1\{y < t\}\right) - 1\left\{\sqrt{2}\cdot\min(x,y) < t\right\}$$

and

$$\Phi_2(x,y;t) = \frac{1}{2}\left(1\left\{\frac{x}{y} < t\right\} + 1\left\{\frac{y}{x} < t\right\}\right) - \frac{t^2}{1+t^2}$$

correspondingly. The projections for every kernel were calculated as

$$\Psi_1(s;t) = \mathbf{E}\left(\Phi_1(X,Y;t)|Y=s\right) = \frac{1}{2}\left(1\{s<t\} - e^{-\frac{t^2}{2}} - 1\right) + 1\left\{s > \frac{t}{\sqrt{2}}\right\}e^{-\frac{t^2}{4}},$$

$$\Psi_2(s;t) = \mathbf{E}\left(\Phi_2(X,Y;t)|Y=s\right) = \frac{1}{2}\left(e^{-\frac{s^2}{2t^2}} + 1 - e^{-\frac{t^2 s^2}{2}}\right) - \frac{t^2}{1+t^2}.$$

The variance function of this projection and its supremum were found as

$$\Delta^2_{i,KU} = \sup_{t\geq 0}\Delta^2_{i,KU}(t) = \sup_{t\geq 0}\mathbf{E}\,\Psi_i^2(X;t), \quad i = 1,2;$$

$$i=1, \quad \Delta^2_{1,KU} = \sup_{t\geq 0}\left(\frac{1}{4}e^{-t^2}\left(e^{\frac{t^2}{2}} - 1\right)\right) = \frac{1}{16}; \text{ at point } t = \sqrt{2\ln(2)};$$

$$i=2, \quad \Delta^2_{2,KU} = \sup_{t\geq 0}\left(\frac{(t-1)^2\, t^2\, (t+1)^2\, (t^4+3t^2+1)}{4\,(t^2+1)^2\,(t^2+2)\,(t^2-t+1)\,(t^2+t+1)\,(2t^2+1)}\right)$$
$$= 0.00954; \text{ at points } t = 0.445 \quad \text{and} \quad t = 2.257.$$

From the calculations in chapter 5 it follows that the kernels are non-degenerate. Since the family of kernels is centered and bounded, we used the theorem from [21] on the logarithmic asymptotics of the probabilities of the large deviation under the null hypothesis H_0, for the U-empirical Kolmogorov-type statistics to the following relation to formulate the next result.

Theorem 13. *For any $z > 0$*

$$\lim_{n\to\infty} n^{-1}\ln\mathbf{P}\left\{KU_{i,n} > z\right\} = h_i(z) \sim -\frac{z^2}{8\Delta^2_{i,KU}} = -w_i(z), \quad i = 1,2,$$

where w_i is a continuous function for which the asymptotic near zero is essential.

Now we describe the alternatives $f_i(x,\theta)$, $x \geq 0$, $i = 1,\ldots,4$, which are considered below.

1. Weibull alternative with density

$$f_1(x,\theta) = \frac{(1+\theta)}{2^\theta}\, x^{2\theta+1}\, \exp\left(-\frac{x^{2\,(1+\theta)}}{2^{1+\theta}}\right);$$

2. Lehmann alternative with density

$$f_2(x,\theta) = (1+\theta)\, f(x)\, F^\theta(x) = (1+\theta) x\, e^{-\frac{x^2}{2}} \left(1 - e^{-\frac{x^2}{2}}\right)^\theta;$$

3. Gamma alternative with density

$$f_3(x,\theta) = \frac{x^{\theta+1}\, e^{-\frac{x^2}{2}}}{2^{\frac{\theta}{2}}\, \Gamma\left(\frac{\theta}{2}+1\right)};$$

4. Rice alternative with density

$$f_4(x,\theta) = x\, \exp\left(-\frac{(x^2+\theta^2)}{2}\right)\, I_0(x\cdot\theta),$$

where $I_0(\cdot)$ is the modified Bessel function of the first kind of order 0.

The Kullback-Leibler information was calculated in chapter 5 of the dissertation research. Since in our case the null hypothesis H_0 is composite, for alternative density $f(x,\theta)$ $K(\theta)$ is defined as follows:

$$K(\theta) = \inf_{\sigma>0} \int_0^\infty \ln \frac{f(x,\theta)}{r(x,\sigma)}\, f(x,\theta)\, dx,$$

where $r(x,\sigma) = \frac{x}{\sigma^2}\, \exp(-\frac{x^2}{2\sigma^2})$.

Lemma. *For some alternative density $f(x,\theta)$ as $\theta \to 0$ the following expression for the Kullback-Leibler information can be obtained*

$$2K(\theta) = \theta^2 \left(I(0) - \left(\int_0^\infty \left(\frac{x}{\sqrt{2}}\right)^2 f'_\theta(x,0)\, dx \right)^2 \right),$$

where I(0) is the Fisher information.

However, this asymptotic θ^2 is not sufficient for the Rice alternative, because $f'_{4,\theta}(x,0) \equiv 0$. In the case of this alternative the following was obtained $K_4(\theta) = \frac{1}{128}\,\theta^8 + o(\theta^8)$. The values of the Kullback-Leibler information found in this chapter of dissertation research are collected in the table 9.

We collect all obtained values of the local Bahadur efficiency for the test statistics in the table 10. We placed the maximum attainable value of the local Bahadur efficiency under $IU^R_{2,\sigma}$.

Table 9. The as $\theta \to 0$

	f_1	f_2	f_3	f_4
$2K(\theta)$	$\frac{\pi^2}{6}\,\theta^2$	$\left(\frac{\pi^2}{3} - \frac{\pi^4}{36}\right)\,\theta^2$	$\left(\frac{\pi^2}{6} - 1\right)\,\theta^2$	$\frac{1}{128}\,\theta^8$

Table 10. The Local Bahadur efficiency for the test statistics

	Alternatives			
	f_1	f_2	f_3	f_4
IU_1^R	0.697	0.807	0.198	0.149
IU_2^R	0.802	0.805	0.202	0.288
$IU_{2,\sigma}^R$	**0.825**	**0.938**	0.23	0.314
KU_1^R	0.158	0.181	0.043	**0.697**
KU_2^R	0.798	0.886	**0.875**	0.243
$L^R(1)$	0.735	0.935	0.227	0.151
$L^R(2)$	0.725	0.931	0.216	0.146
$L^R(3)$	0.711	0.923	0.224	0.138

At the end of the chapter we considered the alternative distributions from [9] and the Rice distribution for computing empirical power for the statistics KU_2^R, $IU_2^R(\sigma), \sigma = 1, 2, 4, 8$ and $L^R(a), a = 1, 2, 5$ for the sample size equal to 20 and 0.1, 0.05 and 0.01 significance levels.

For ease of perception of the results presented in tables 11 and 12, the following notations are given: W – weibull, G – gamma, IG – inverse Gauss, LN – log normal, GO – Gompertz law, PL – Power law, LFL – linear failure law, EP – exponential power law, PE – Poisson-exponential, RL – Rice law, Rice(1,t), scale-parameter equals 1.

Table 11. Power for the$IU_2^R(\cdot)$ and KU_2^R for the 0.1, 0.05 and 0.01 significance levels

	IU_2^R, $\sigma=1$			IU_2^R, $\sigma=2$			IU_2^R, $\sigma=4$			IU_2^R, $\sigma=8$			KU_2^R		
alternatives	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
Exp	94	91	81	96	94	87	96	93	85	94	90	78	96	92	80
W(2,1)	10	5	1	10	5	1	10	5	1	10	5	1	10	5	1
W(3,1)	60	46	22	62	48	22	57	44	21	51	39	18	64	51	28
G(1.5,1)	73	64	44	75	66	45	73	63	42	63	51	30	71	59	33
G(2,1)	44	33	16	43	31	14	38	27	12	28	18	6	27	25	8
IG(1,0.5)	94	91	80	95	92	81	94	90	77	87	81	62	94	89	72
IG(1,1)	70	60	38	66	54	30	60	47	24	40	26	9	62	47	20
IG(1,1.5)	38	26	10	30	19	6	23	13	3	9	4	1	27	16	3
IG(2,0.5)	99	98	94	99	99	96	99	99	96	98	97	92	99	98	93
IG(2,1)	94	91	80	95	92	81	94	90	77	89	81	61	94	90	73
IG(2,1.5)	85	77	59	84	76	54	81	71	48	66	52	28	80	70	43
LN(0,0.8)	57	45	25	52	40	19	45	33	14	28	17	5	47	33	12
LN(0,1.2)	97	96	89	98	97	91	98	96	90	96	92	81	97	95	86
LN(0,1.5)	*	99	98	*	*	99	*	*	99	*	99	98	*	*	98
GO(0.5)	80	72	52	85	79	63	85	79	62	82	74	56	83	74	51
GO(1)	62	52	32	71	62	43	73	64	44	71	61	42	69	57	35
GO(1.5)	48	38	20	59	49	30	62	52	33	62	52	33	56	44	23
PL(1)	27	18	8	43	33	18	52	41	24	57	47	29	49	37	18
PL(1.5)	77	70	53	91	86	75	93	90	80	94	90	81	91	86	71
PL(2)	95	92	84	99	98	95	99	99	97	99	99	97	99	98	94
LFL(1)	76	67	47	82	74	57	82	74	56	78	69	51	78	68	45
LFL(2)	64	54	33	71	62	43	72	62	43	68	58	39	67	56	32
LFL(3)	56	46	26	64	53	34	65	54	35	62	51	33	60	47	25
LFL(4)	50	40	22	58	48	29	59	49	29	57	47	28	54	42	21
EP(0.5)	*	99	98	*	*	*	*	*	*	*	*	99	*	*	99
EP(1)	62	52	32	71	62	42	73	64	44	71	61	42	68	57	34
EP(2)	34	23	8	30	20	8	26	17	6	24	16	6	34	25	11
EP(3)	92	85	62	90	83	61	81	72	50	71	62	43	92	86	68
PE(1)	92	88	77	95	92	82	95	91	81	92	87	73	93	88	73
PE(2)	85	78	62	89	83	68	88	82	66	84	76	59	86	78	57
PE(3)	68	58	40	73	64	45	73	64	44	68	58	38	69	58	33
PE(4)	46	36	19	51	40	23	51	40	22	47	36	19	45	33	15
RL(2)	40	27	10	38	26	10	35	24	10	31	22	8	42	31	14
RL(3)	91	83	59	91	83	59	86	78	55	78	69	48	92	86	67
RL(4)	*	*	96	*	*	96	99	98	93	97	95	87	*	*	98

Table 12. Power for the L(20,s) for the 0.1, 0.05 и 0.01 significance levels

	L, $a=1$			L, $a=2$			L, $a=5$		
Alternatives	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
Exp	96	93	84	96	93	87	95	92	81
W(2,1)	10	5	1	10	5	1	10	5	1
W(3,1)	59	46	22	58	44	20	58	45	21
G(1.5,1)	71	60	38	71	60	38	68	57	34
G(2,1)	36	25	9	36	25	10	32	21	8
IG(1,0.5)	93	88	72	93	88	72	91	85	66
IG(1,1)	55	41	17	54	40	17	48	34	12
IG(1,1.5)	18	9	2	18	9	1	13	6	1
IG(2,0.5)	99	98	95	99	98	95	99	98	93
IG(2,1)	93	88	72	93	88	71	91	85	66
IG(2,1.5)	78	66	41	76	65	39	73	60	33
LN(0,0.8)	40	27	10	39	27	10	34	22	6
LN(0,1.2)	97	95	87	98	95	87	97	94	84
LN(0,1.5)	*	*	99	*	*	99	*	*	98
GO(0.5)	85	78	62	85	78	62	84	77	60
GO(1)	74	65	46	74	65	46	73	64	45
GO(1.5)	64	54	34	64	55	36	64	54	36
PL(1)	55	45	27	56	46	29	57	47	29
PL(1.5)	94	91	83	95	92	83	94	91	83
PL(2)	99	99	97	99	99	98	99	98	97
LFL(1)	82	75	57	82	74	56	80	73	54
LFL(2)	72	63	43	72	63	44	71	62	42
LFL(3)	66	56	36	65	55	36	66	55	36
LFL(4)	61	51	31	61	51	32	60	50	31
EP(0.5)	*	*	*	*	*	*	*	*	*
EP(1)	74	65	46	74	66	47	74	65	46
EP(2)	27	18	7	27	18	7	26	18	7
EP(3)	83	75	53	83	74	53	82	73	52
PE(1)	94	91	80	94	91	80	93	89	77
PE(2)	88	82	66	87	82	66	86	80	63
PE(3)	73	63	44	73	64	44	71	62	42
PE(4)	52	41	23	51	41	23	50	40	22
RL(2)	36	25	10	36	25	10	35	24	9
RL(3)	87	79	57	87	79	56	86	78	56
RL(4)	*	99	94	99	99	94	99	99	94

In the conclusion section the main results of this research were formulated and listed.

References

[1] Baringhaus L., Henze N. A class of consistent tests for exponentiality based on the empirical Laplace transform. Annals of the Institute of Statistical Mathematics. 1991;43(3):551–564.

[2] Baringhaus L., Henze N. Tests of fit for exponentiality based on a characterization via the mean residual life function. Statistical Papers. 2000;41:225–236.

[3] Nikitin, Y.Y. *Tests based on characterizations, and their efficiencies: a survey.* Acta et Commentationes Universitatis Tartuensis de Mathematica, 21(2017), N 1, 3-24.

[4] Ahsanullah M, Anis MZ. Some characterizations of exponential distribution. International Journal of Statistics and Probability. 2017; 6(5):132–139

[5] Hoeffding W. *A class of statistics with asymptotically normal distribution.* Ann. Math. Statist.– 1948. V. 19.–P. 293-325.

[6] M. Ahsanullah, V.B. Nevzorov, G.P. Yanev, *Characterizations of distributions via order statistics with random exponential shifts*, J. Appl. Statist. Sci., 18(2010), pp. 297-–305.

[7] V. S. Korolyuk, Yu. V. Borovskikh, *Theory of U-statistics.*, Kluwer, Dordrecht, 1994

[8] Chin-Yuan Hua, Gwo Dong Lin.*Characterizations of the logistic and related distributions*. Journ. of Mathem. Anal. and Appl, 463(2018), N 1, 79–92

[9] Meintanis S., Iliopoulos, G. *Tests of fit for the Rayleigh distribution based on the empirical Laplace transform.* Ann Inst Stat Math 55, 137–151 (2003). https://doi.org/10.1007/BF02530490

[10] Muliere P., Nikitin Ya., Yu. Scale-invariant test of normality based on Polya's characterization. Metron. V. 60 (1-2), pp. 21-33, 2002.

[11] Ya. Yu. Nikitin, K. Yu. Volkova, *Efficiency of Exponentiality Tests Based on a Special Property of Exponential Distribution,* Math. Meth. Stat., 25(2016), pp. 54–66.

[12] Bahadur R.R. Some limit theorems in statistics. SIAM, Philadelphia (1971).

[13] Bahadur R. R. *Stochastic comparison of tests* Ann. Math. Stat. 31(1960), 276–295.

[14] Linnik YU. V. Lineynyye formy i statisticheskiye kriterii. I,II.// Ukrainskiy matematicheskiy zhurnal, 1953, 5 (2), 207-243; 5 (3), 247-290.

[15] Ya. Nikitin, *Asymptotic efficiency of nonparametric tests,* Cambridge University Press, New York, 1995.

[16] Nikitin YY, Tchirina AV. *Bahadur efficiency and local optimality of a test for the exponential distribution based on the Gini statistic*. Statistical Methods & Applications. 1996;5(1):163–175.

[17] Nikitin Y. Y., Tchirina A. V. Lilliefors test for exponentiality: Large deviations, asymptotic efficiency, and conditions of local optimality. Mathematical Methods of Statistics. 2007;16(1):16–24.

[18] M. Obradović, M. Jovanovic, B. Milošević, *Goodness-of-fit tests for Pareto distribution based on a characterization and their asymptotics.* — J. Theor. Appl. Statist. **49**, No. 5 (2015), 1026–1041.

[19] Nikitin Ya. Yu., Ponikarov E.V. *Rough large deviation asymptotics of Chernoff type for von Mises functionals and U-statistics.* Proc. of St.Petersburg Math. Soc., **7**(1999), 124–167. Engl. transl. in AMS Transl., ser.2, **203**(2001), 107–146.

[20] Nikitin Ya. Yu., Peaucelle I. *Efficiency and local optimality of distribution-free tests based on U- and V- statistics.* Metron, **LXII**(2004),185–200.

[21] Nikitin Ya. Yu. *Large deviations of U-empirical Kolmogorov-Smirnov tests, and their efficiency.* J. Nonpar. Stat., 22(2010), 649–668.

[22] Nikitin Y. Y., Volkova K. Y. Exponentiality tests based on Ahsanullah's characterization and their efficiency. Journ of Mathem Sciences. 2015;204(1):42–54.

[23] A. Asgharzadeh, R. Valiollahi, M. Abdi, *Point and interval estimation for the logistic distribution based on record data.* SORT-Stat. Oper. Res. Trans., 1(2016), pp. 89-112.

[24] Volkova K. YU. Asimptoticheskaya effektivnost' kriteriyev soglasiya, osnovannykh na kharakterizatsionnykh svoystvakh raspredeleniy. Kandidatskaya dissertatsiya. SPbGU. 2011.

[25] Litvinova V.V. Asimptoticheskiye svoystva kriteriyev simmetrii i soglasiya, osnovannykh na kharakterizatsiyakh. Kandidatskaya dissertatsiya. SPbGU. 2004.

[26] Litvinova V. V., Nikitin YA.YU. Dva semeystva kriteriyev normal'nosti, osnovannykh na kharakterizatsii Poya, i ikh asimptoticheskaya effektivnost'. Zap. nauchn. semin. POMI, 328, 147-159, 2005.

[27] Chirina A. V. Asimptoticheskaya effektivnost' i lokal'naya optimal'nost' po Bakhaduru kriteriya eksponentsial'nosti, osnovannogo na statistike Morana. Zap. nauchn. semin. POMI, 294, 245-259, 2002.

[28] Chirina A. V. Asimpototicheskaya effektivnost' kriteriyev eksponentsial'nosti, svobodnykh ot parametra masshtaba. Kandidatskaya dissertatsiya. SPbGU. 2005.

Authors publicacions.

[29] Nikitin YA. YU., Ragozin I. A. *Kriterii soglasiya, osnovannyye na kharakterizatsii logisticheskogo raspredeleniya.* Vestnik SPbGU. Matematika. Mekhanika. Astronomiya, 2019, 6 (64), 241-252.

[30] Y. Y. Nikitin, I. A. Ragozin, *Goodness-of-Fit Tests Based on a Characterization of Logistic Distribution.* Vestnik St. Petersburg Univ., Mathematics, 52(2019), pp. 169-177.

[31] Ya. Yu. Nikitin, I. A. Ragozin. *Goodness-of-fit tests for the logistic location family.* Journal of Applied Statistics, 2020, 47:13-15, 2610-2622, DOI: 10.1080/02664763.2020.1761952

[32] J. S. Allison, Ya. Yu. Nikitin, I. A. Ragozin, L. Santana. *New tests for exponentiality based on a characterization with random shift.* Journal of Statistical Computation and Simulation, 2020, 90:15, 2840-2857, DOI: 10.1080/00949655.2020.1791865

[33] Ragozin I. A. *Novyye kriterii soglasiya dlya raspredeleniya Pareto I tipa, osnovannyye na nekotoroy kharakterizatsii.* Zapiski nauchnykh seminarov POMI RAN, 2020, 495, 237-249.

[34] Ragozin I. A. *Novyye kriterii soglasiya dlya semeystva raspredeleniy Releya, osnovannyye na nekotorom spetsial'nom svoystve i nekotoroy kharakterizatsii.* Zapiski nauchnykh seminarov POMI RAN, 2021, 505, 230-243.

[35] Ragozin I. A., *New goodness-of-fit tests for Pareto I type distribution, based on some characterization.* Journal of Mathematical Sciences, 2022, Vol. 268, No.5, P.684-692, https://doi.org/10.1007/s10958-022-06238-4

[36] Ragozin I. A., *New goodness-of-fit tests for family of Rayleigh distributions, based on a special property and a characterization.* Journal of Mathematical Sciences, Vol. 281, No. 1, 2024. DOI 10.1007/s10958-024-07083-3