

Goodness-of-fit tests for the Pareto distribution based on its characterization

Ksenia Volkova¹

Accepted: 7 July 2015 / Published online: 25 July 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract A new characterization of the Pareto distribution is proposed, and new goodness-of-fit tests based on it are constructed. Test statistics are functionals of U -empirical processes. The first of these statistics is of integral type, it is similar to the classical statistics ω_n^1 . The second one is a Kolmogorov type statistic. We show that the kernels of our statistics are non-degenerate. The limiting distribution and large deviations asymptotics of the new statistics under null hypothesis are described. Their local Bahadur efficiency for parametric alternatives is calculated. This type of efficiency is mostly appropriate for the solution of our problem since the Kolmogorov type statistic is not asymptotically normal, and the Pitman approach is not applicable to this statistic. For the second statistic we evaluate the critical values by using Monte-Carlo methods. Also conditions of local optimality of new statistics in the sense of Bahadur are discussed and examples of such special alternatives are given. For small sample size we compare the power of those tests with some common goodness-of-fit tests.

Keywords Pareto distribution · U -statistics · Characterization · Bahadur efficiency · Goodness-of-fit test

Mathematics Subject Classification 60F10 · 62G10 · 62G20 · 62G30

Research supported by Grant RFBR No. 13-01-00172, Grant NSh No. 2504.2014.1 and by SPbGU Grant No. 6.38.672.2013.

✉ Ksenia Volkova
efrksenia@gmail.com

¹ Saint-Petersburg State University, Universitetsky pr. 28, Stary Peterhof 198504, Russia

1 Introduction

Let \mathcal{P} be the family of the Pareto distributions with the distribution function (d.f.)

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

In this paper we develop the goodness-of-fit tests for the Pareto distribution using a new characterization based on the property of order statistics. The problem formulation is as follows: let X_1, \dots, X_n be positive i.i.d. random variables (rv's) with continuous d.f. F . Consider testing the composite hypothesis $H_0 : F \in \mathcal{P}$ against the general alternative $H_1 : F \notin \mathcal{P}$, assuming that the alternative d.f. is also concentrated on $[1, \infty)$.

The goodness-of-fit tests for the Pareto distribution have been discussed in [Beirlant et al. \(2006\)](#), [Gulati and Shapiro \(2008\)](#), [Martynov \(2009\)](#), [Rizzo \(2009\)](#). We exploit a different idea for constructing and analyzing statistical tests based on characterization by the property of equidistribution of linear statistics by means of so-called U -empirical d.f.'s ([Janssen 1988](#); [Korolyuk and Borovskikh 1994](#)). This method was developed early in several articles, particularly, in [Nikitin \(1996\)](#), [Nikitin and Peacelle \(2004\)](#), [Nikitin and Tchirina \(1996\)](#), [Nikitin and Volkova \(2010\)](#), [Nikitin and Volkova \(2012\)](#), [Litvinova \(2004\)](#). The tests for the Pareto distribution using this approach were obtained and analyzed in [Jovanovic et al. \(2014\)](#). One can observe that the new tests based on characterizations have reasonably high efficiencies and can be competitive with previously known goodness-of-fit tests. Let us explain our approach.

We will say that the d.f. F belongs to the class of distributions \mathcal{F} , if $\forall x_1, x_2$: either $F(x_1x_2) \leq F(x_1)F(x_2)$ or $F(x_1x_2) \geq F(x_1)F(x_2)$, see [Ahsanullah \(1989\)](#).

Let X_1, \dots, X_n be i.i.d. positive absolutely continuous random variables with the d.f. F from the class \mathcal{F} . Denote by $X_{(1,n)} \leq X_{(2,n)} \leq \dots \leq X_{(n,n)}$ - the order statistics of a random sample X_1, \dots, X_n .

We present a new characterization within the class \mathcal{F} .

Theorem 1 *For the fixed k let X_1, \dots, X_k be i.i.d., positive and bounded rv's having an absolutely continuous (with respect to Lebesgue measure) d.f. $F(x)$. Then the equality in law of X_1 and $X_{(k,k)}/X_{(k-1,k)}$ takes place iff X_1 has some d.f. from the family \mathcal{P} .*

Proof Let $Y = \ln X$ and let G denote the d.f. of Y . It can be easily seen that $F \in \mathcal{F}$ iff G is NBU ("new better than used") or NWU ("new worse than used") ([Ahsanullah 1977](#)). Further, since we use the monotonic transformation, then X_1 and $X_{(k,k)}/X_{(k-1,k)}$ will be identically distributed iff Y_1 and $Y_{(k,k)} - Y_{(k-1,k)}$ are identically distributed. It follows from [Ahsanullah \(1977\)](#) that X_1 and $X_{(k,k)}/X_{(k-1,k)}$ are identically distributed iff $Y = \ln X$ has the exponential distribution with some scale parameter λ , therefore X_1 has the Pareto distribution with the same parameter λ . \square

In the case when $k = 2$ our characterization coincide with another characterization of the Pareto distribution considered in [Jovanovic et al. \(2014\)](#), see also [Nikitin and Volkova \(2012\)](#). Note that our characterization extends the charaterization, involved in [Jovanovic et al. \(2014\)](#).

According to our characterization we construct the U -empirical d.f. by the formula

$$H_n(t) = \binom{n}{k}^{-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} \mathbf{1}\{X_{(k, \{i_1, \dots, i_k\})} / X_{(k-1, \{i_1, \dots, i_k\})} < t\}, \quad t \geq 1,$$

where $X_{(s, \{i_1, \dots, i_k\})}$, $s = \{k - 1, k\}$ denotes the s -th order statistic of the subsample X_{i_1}, \dots, X_{i_k} . For rv X the U -statistical d.f. will be simply the usual empirical d.f. $F_n(t) = n^{-1} \sum_{i=1}^n \mathbf{1}(X_i < t)$, $t \in \mathbb{R}^1$, based at the observations X_1, \dots, X_n .

It is known that the properties of U -empirical d.f.'s are similar to the properties of usual empirical d.f.'s (Helmers et al. 1988; Janssen 1988). Hence the difference $H_n - F_n$ for large n should be almost surely close to zero under H_0 , and we can measure their closeness by using some test statistics, assuming their large values to be critical.

We suggest two test statistics

$$I_n^{(k)} = \int_1^\infty (H_n(t) - F_n(t)) dF_n(t), \tag{2}$$

$$D_n^{(k)} = \sup_{t \geq 1} |H_n(t) - F_n(t)|. \tag{3}$$

Note that both proposed statistics under H_0 are invariant with respect to the change of variables $X \rightarrow X^{\frac{1}{\lambda}}$, so we may set $\lambda = 1$.

We discuss their limiting distributions under the null hypothesis and find logarithmic asymptotics of large deviations under H_0 . Next we calculate their efficiencies against some parametric alternatives from the class \mathcal{F} . We use the notion of local exact Bahadur efficiency (BE) (Bahadur 1971; Nikitin 1995), as the statistics $D_n^{(k)}$ has the non-normal limiting distribution, hence the Pitman approach to the efficiency is not applicable. However, it is known that the local BE and the limiting Pitman efficiency usually coincide, see Wieand (1976), Nikitin (1995).

Finally, we study the conditions of the local optimality of our tests, describe the "most favorable" alternatives for them and compare the powers of our tests with some standard goodness-of-fit tests.

The family of d.f. in null-hypothesis we consider is a particular case of the so-called Pareto type I distribution with the d.f. $P_1(x) = 1 - (\frac{x}{\beta})^{-\lambda}$, $x \geq \beta > 0$, $\lambda > 0$, see, for example, Arnold (1983). For practice goodness-of-fit testing based on our new tests the unknown parameters of the hypothesized Pareto distribution can be estimated by a number of methods, see Arnold (1983, Ch. 5), Kleiber and Kotz (2003, Ch. 3), Brazauskas and Serfling (2003), Rizzo (2009). One can estimate first the parameter β for example by the MLE estimator $\hat{\beta} = \min_{i=1, \dots, n} X_i$. Then the sample X_1, \dots, X_n can be transformed to the new sample Y_1, \dots, Y_n , where $Y_i = X_i / \hat{\beta}$ and its has the d.f. considered in (1).

There exist the Pareto's second model, so-called Pareto type II distribution with d.f. $P_2(x) = 1 - (1 + \frac{x-\mu}{\beta})^{-\lambda}$, $x \geq \mu \in \mathbb{R}$, $\lambda > 0$. Pareto type I and type II models are related by a following transformation: if rv X has a Pareto type II distribution, then $X - (\mu - \sigma)$ has a Pareto type I distribution. Therefore using one of the estimator

of the location parameter μ , one can reduce Pareto type II rv to our model. We do not discuss the difference between the parameter estimators and concentrate on the constructing of the goodness-of-fit tests.

2 Integral statistic $I_n^{(k)}$

The statistic $I_n^{(k)}$ is asymptotically equivalent to the U -statistic of degree $(k + 1)$ with the centered kernel

$$\psi_k(X_{i_1}, \dots, X_{i_{k+1}}) = \frac{1}{k + 1} \sum_{\pi(i_1, \dots, i_{k+1})} \mathbf{1}(X_{(k, \{i_1, \dots, i_k\})} / X_{(k-1, \{i_1, \dots, i_k\})} < X_{i_{k+1}}) - \frac{1}{2},$$

where $\pi(i_1, \dots, i_{k+1})$ means all permutations of different indices from $\{i_1, \dots, i_{k+1}\}$.

Let X_1, \dots, X_{k+1} be independent rv's from the standard Pareto distribution. It is known that non-degenerate U -statistics are asymptotically normal (Hoeffding 1948; Korolyuk and Borovskikh 1994). To prove that the kernel $\Psi_k(X_1, \dots, X_{k+1})$ is non-degenerate, we calculate its projection $\psi_k(s)$. For a fixed $X_{k+1} = s, s \geq 1$ we have:

$$\begin{aligned} \psi_k(s) &:= E(\Psi_k(X_1, \dots, X_{k+1}) \mid X_{k+1} = s) \\ &= \frac{k}{k + 1} \mathbb{P}(X_{(k, \{2, \dots, k, s\})} / X_{(k-1, \{2, \dots, k, s\})} < X_1) \\ &\quad + \frac{1}{k + 1} \mathbb{P}(X_{(k, \{1, \dots, k\})} / X_{(k-1, \{1, \dots, k\})} < s) - \frac{1}{2}. \end{aligned}$$

It follows from the above characterization that the second probability is equal to:

$$\mathbb{P}(X_{k, \{1, \dots, k\}} / X_{k-1, \{1, \dots, k\}} < s) = \mathbb{P}(X_1 < s) = F(s).$$

It remains to calculate the first term. For this purpose we decompose the probability as $\mathbb{P}(X_{k, \{2, \dots, k, s\}} / X_{k-1, \{2, \dots, k, s\}} < X_1) = \mathbb{P}_1 + \mathbb{P}_2 + \mathbb{P}_3$, where $\mathbb{P}_i, i = 1, 2, 3$ are initial probabilities, computed in one of the following cases:

- (1) Let the sample units take places as follows: $X_2 < \dots < X_k < s$. Then our probability transforms into

$$\begin{aligned} \mathbb{P}_1 &= (k - 1)! \mathbb{P}\left(\frac{s}{X_k} < X_1, X_2 < \dots < X_k < s\right) \\ &= (k - 1)! \mathbb{P}\left(X_k < s, X_1 > \frac{s}{X_k}, X_2 < X_3, X_3 < X_4, \dots, X_{k-1} < X_k\right). \end{aligned}$$

After some calculations we obtain that the last probability is equal to:

$$\begin{aligned} & (k - 1)! \int_1^s \left(1 - F \left(\frac{s}{x_k} \right) \right) \frac{F^{k-2}(x_k)}{(k - 2)!} dF(x_k) \\ & = F^{k-1}(s) - (k - 1) \int_1^s \left(1 - \frac{1}{x} \right)^{k-2} \left(1 - \frac{x}{s} \right) \frac{dx}{x^2}. \end{aligned}$$

The integral in the second term can be evaluated using integration by parts and binomial representation of the function $\left(1 - \frac{1}{x} \right)^{k-1}$. Finally we have:

$$\begin{aligned} & \int_1^s \left(1 - \frac{1}{x} \right)^{k-2} \left(1 - \frac{x}{s} \right) \frac{dx}{x^2} = \frac{1}{s(k - 1)} \int_1^s \sum_{j=0}^{k-1} (-1)^j \binom{k - 1}{j} x^{-j} dx \\ & = \frac{1}{s(k - 1)} \left(s - 1 - (k - 1) \ln(s) + \sum_{j=2}^{k-1} (-1)^j \binom{k - 1}{j} \frac{1 - s^{-(j-1)}}{j - 1} \right). \end{aligned}$$

Thus the initial probability in this case is equal to

$$\mathbb{P}_1 = F^{k-1}(s) - F(s) + (k - 1) \frac{\ln s}{s} - \frac{1}{s} \sum_{j=2}^{k-1} (-1)^j \binom{k - 1}{j} \frac{1 - s^{-(j-1)}}{j - 1}.$$

(2) The sample units are $X_2 < X_3 < \dots < X_{k-1} < s < X_k$, then for this case we have:

$$\begin{aligned} \mathbb{P}_2 & = (k - 1)! \mathbb{P} \left(\frac{X_k}{s} < X_1, X_2 < X_3 < \dots < X_{k-1} < s < X_k \right) \\ & = (k - 1)! \mathbb{P} \left(X_k > s, X_1 > \frac{X_k}{s}, X_2 < X_3, X_3 < X_4, \dots, X_{k-1} < s \right) \\ & = (k - 1)! \int_s^\infty \left(1 - F \left(\frac{x_k}{s} \right) \right) \frac{F^{k-2}(s)}{(k - 2)!} dF(x_k) \\ & = \frac{(k - 1)}{2s} F^{k-2}(s). \end{aligned}$$

(3) The last case we consider is when s is situated on j -th place ($1 \leq j \leq k - 2$) in variational series of the sample X_2, \dots, X_{k-2} . It means that the sample units take places as follows: $X_2 < \dots < s < \dots < X_{k-2} < X_{k-1} < X_k$ and s also may stand on first and $(k - 2)$ -th places. Then the required probability is equal to

$$\begin{aligned} \mathbb{P}_3 & = (k - 1)! \mathbb{P} \left(\frac{X_k}{X_{k-1}} < X_1, X_2 < \dots < s < \dots < X_{k-2} < X_{k-1} < X_k \right) \\ & = \frac{1}{2} C_{k-1}^{j-1} (1 - F(s))^{k-j} F^{j-1}(s), \quad 1 \leq j \leq k - 2. \end{aligned}$$

Combining the results we get that the first term in the projection has the form:

$$\mathbb{P}(X_{(k, \{2, \dots, k, s\})} / X_{(k-1, \{2, \dots, k, s\})} < X_1) = F^{k-1}(s) - F(s) + (k - 1) \frac{\ln s}{s} - \frac{1}{s} \sum_{j=2}^{k-1} (-1)^j \binom{k-1}{j} \frac{1 - s^{-(j-1)}}{j-1} + \frac{1}{2} \sum_{j=1}^{k-1} C_{k-1}^{j-1} (1 - F(s))^{k-j} F^{j-1}(s).$$

Note that the last sum is equal to $\sum_{j=1}^{k-1} C_{k-1}^{j-1} (1 - F(s))^{k-j} F^{j-1}(s) = 1 - F^{k-1}(s)$. Thus for the initial probability we get the result:

$$\mathbb{P}(X_{(k, \{2, \dots, k, s\})} / X_{(k-1, \{2, \dots, k, s\})} < X_1) = \frac{1}{2} F^{k-1}(s) - F(s) + (k - 1) \frac{\ln s}{s} - \frac{1}{s} \sum_{j=2}^{k-1} (-1)^j \binom{k-1}{j} \frac{1 - s^{-(j-1)}}{j-1} + \frac{1}{2}.$$

Hence we get the final expression for the projection of the kernel ψ_k :

$$\psi_k(s) = \frac{kF^{k-1}(s) - 1}{2(k + 1)} - \frac{k - 1}{k + 1} F(s) + \frac{k(k - 1) \ln s}{k + 1} \frac{1}{s} - \frac{k}{s(k + 1)} \sum_{j=2}^{k-1} (-1)^j \binom{k-1}{j} \frac{1 - s^{-(j-1)}}{j-1}. \tag{4}$$

The calculation of the variance for the projection ψ_k in the general case is too complicated, therefore we calculate it only for particular k .

2.1 Integral statistic $I_n^{(3)}$

The projection $\psi_k(s)$ for the case $k = 3$ has the form:

$$\psi_3(s) = \frac{9}{8s^2} + \frac{3 \ln s}{2s} - \frac{1}{s} - \frac{1}{4}. \tag{5}$$

The variance of this projection $\Delta_3^2 = E\psi_3^2(X_1)$ under H_0 is given by

$$\Delta_3^2 = \int_1^\infty \psi_3^2(s) \frac{1}{s^2} ds = \frac{11}{1920} \approx 0.006.$$

Therefore the kernel Ψ_3 is centered and non-degenerate. We can apply Hoeffding’s theorem on asymptotic normality of U -statistics, see again [Hoeffding \(1948\)](#), [Korolyuk and Borovskikh \(1994\)](#), which implies that the following result holds

Theorem 2 *Under the null hypothesis as $n \rightarrow \infty$ the statistic $\sqrt{n}I_n^{(3)}$ is asymptotically normal so that*

$$\sqrt{n}I_n^{(3)} \xrightarrow{d} \mathcal{N}\left(0, \frac{11}{120}\right).$$

Now we shall evaluate the large deviation asymptotics of the sequence of statistics $I_n^{(3)}$ under H_0 . According to the theorem on large deviations of such statistics from [Nikitin and Ponikarov \(1999\)](#), see also [DasGupta \(2008\)](#), [Nikitin \(2010\)](#), we obtain due the fact that the kernel Ψ_3 is centered, bounded and non-degenerate the following result.

Theorem 3 For $a > 0$

$$\lim_{n \rightarrow \infty} n^{-1} \ln P(I_n^{(3)} > a) = -f_I^{(3)}(a),$$

where the function $f_I^{(3)}$ is continuous for the sufficiently small $a > 0$, and

$$f_I^{(3)}(a) \sim \frac{a^2}{32\Delta_3^2} = 5.455 a^2, \quad \text{as } a \rightarrow 0.$$

2.2 Some notions from the Bahadur theory

Suppose that under the alternative H_1 the observations have the d.f. $G(\cdot, \theta)$ and the density $g(\cdot, \theta)$, $\theta \geq 0$, such that $G(\cdot, 0) \in \mathcal{P}$. The measure of the Bahadur efficiency (BE) for any sequence $\{T_n\}$ of test statistics is the exact slope $c_T(\theta)$ describing the rate of an exponential decrease for the attained level under the alternative d.f. $G(\cdot, \theta)$. According to the Bahadur theory ([Bahadur 1971](#); [Nikitin 1995](#)) the exact slopes may be found by using the following Proposition.

Proposition 1 Suppose that the following two conditions hold:

- a) $T_n \xrightarrow{P_\theta} b(\theta)$, $\theta > 0$,
 where $-\infty < b(\theta) < \infty$, and $\xrightarrow{P_\theta}$ denotes convergence in probability under $G(\cdot; \theta)$.
- b) $\lim_{n \rightarrow \infty} n^{-1} \ln P_{H_0}(T_n \geq t) = -h(t)$

for any t in an open interval I , on which h is continuous and $\{b(\theta), \theta > 0\} \subset I$.
 Then

$$c_T(\theta) = 2 h(b(\theta)).$$

We have already found the large deviation asymptotics. In order to evaluate the exact slope it remains to calculate the first condition of the Proposition.

Note that the exact slopes for any θ satisfy the inequality ([Bahadur 1971](#); [Nikitin 1995](#))

$$c_T(\theta) \leq 2K(\theta), \tag{6}$$

where $K(\theta)$ is the Kullback-Leibler “distance” between the alternative and the null-hypothesis H_0 . In our case H_0 is composite, hence for any alternative density $g_j(x, \theta)$ one has

$$K_j(\theta) = \inf_{\lambda > 0} \int_1^\infty \ln[g_j(x, \theta)/\lambda x^{-\lambda-1}]g_j(x, \theta) dx.$$

This quantity can be easily calculated as $\theta \rightarrow 0$ for particular alternatives. According to (6), the local BE of the sequence of statistics T_n is defined as

$$e^B(T) = \lim_{\theta \rightarrow 0} \frac{c_T(\theta)}{2K(\theta)}.$$

2.3 The local Bahadur efficiency of $I_n^{(3)}$

According to Bahadur theory, the considered alternatives should be close to null-hypothesis as $\theta \rightarrow 0$. Therefore we suggest three alternatives against the Pareto distribution. The first two alternatives we consider are obtained by skewing mechanism (Ley and Paindaveine 2008), we call them Ley–Paindaveine alternatives.

i) First Ley–Paindaveine alternative with the d.f.

$$G_1(x, \theta) = F(x)e^{-\theta(1-F(x))}, \quad \theta \geq 0, x \geq 1;$$

ii) Second Ley–Paindaveine alternative with the d.f.

$$G_2(x, \theta) = F(x) - \theta \sin \pi F(x), \quad \theta \in [0, \pi^{-1}], x \geq 1;$$

iii) log-Weibull alternative with the d.f.

$$G_3(x, \theta) = 1 - e^{-(\ln x)^{\theta+1}}, \quad \theta \in (0, 1), x \geq 1.$$

Let us find the local BE for the alternative under consideration.

According to the Law of Large Numbers for U -statistics (Korolyuk and Borovskikh 1994), the limit in probability under H_1 is equal to

$$b_1(\theta) = P_\theta(X_{(3,3)}/X_{(2,3)} < Y) - \frac{1}{2}.$$

It is easy to show (Jovanovic et al. 2014) that

$$b_1(\theta) \sim 4\theta \int_1^\infty \psi_3(s)h_1(s)ds,$$

where $h_1(s) = \frac{\partial}{\partial \theta} g_1(s, \theta) |_{\theta=0}$ and $\psi_3(s)$ is the projection from (5). Therefore for the first Ley–Paindaveine alternative we have

$$b_1(\theta) \sim 4\theta \int_1^\infty \left(\frac{9}{8s^2} + \frac{3 \ln s}{2s} - \frac{1}{s} - \frac{1}{4} \right) \left(\frac{s-2}{s^3} \right) \frac{ds}{s^2} \sim \frac{\theta}{12}, \quad \theta \rightarrow 0,$$

Table 1 Local Bahadur efficiencies for $I_n^{(3)}$

Alternative	Efficiency
Ley-Paindaveine 1	0.909
Ley-Paindaveine 2	0.905
log-Weibull	0.787

and the local exact slope of the sequence $I_n^{(3)}$ as $\theta \rightarrow 0$ admits the representation

$$c_1(\theta) = b_1^2(\theta)/(16\Delta_3^2) \sim \frac{5}{66} \theta^2, \quad \theta \rightarrow 0.$$

The Kullback-Leibler “distance” $K_1(\theta)$ between the alternative and the null-hypothesis H_0 admits the following asymptotics (Jovanovic et al. 2014):

$$2K_1(\theta) \sim \theta^2 \left[\int_1^\infty h_1^2(x)x dx - \left(\int_1^\infty h_1(x) \ln(x) dx \right)^2 \right], \quad \theta \rightarrow 0.$$

Therefore in our case

$$K_1(\theta) \sim \theta^2/24, \quad \theta \rightarrow 0. \tag{7}$$

Consequently, the local efficiency of the test is

$$e_1^B(I) = \lim_{\theta \rightarrow 0} \frac{c_1(\theta)}{2K_1(\theta)} \approx \frac{10}{11} \approx 0.909.$$

Omitting the calculations similar to previous cases, we get for the second Ley-Paindaveine alternative $b_2(\theta) \sim 0.353 \theta$, $c_2(\theta) \sim 1.363 \theta^2$, $\theta \rightarrow 0$. It is easy to show that $K_2(\theta) \sim 0.753 \theta^2$, $\theta \rightarrow 0$. Therefore the local BE is equal to 0.905.

After some calculations in case of the log-Weibull alternative we have:

$$b_3(\theta) \sim \left(\frac{3}{4} - \ln 3 + \ln 2 \right) \theta \approx 0.345 \theta, \quad \theta \rightarrow 0,$$

and the local exact slope of the sequence $I_n^{(3)}$ as $\theta \rightarrow 0$ admits the representation $c_3(\theta) \sim 1.295 \theta^2$. Moreover for the log-Weibull distribution $K_3(\theta)$ satisfies $K_3(\theta) \sim \frac{\theta^2}{12}$, $\theta \rightarrow 0$. Hence the local BE for the last case is equal to 0.787.

Table 1 gathers values of the local BE.

2.4 Integral statistic $I_n^{(4)}$

For the case $k = 4$ the projection $\psi_k(s)$ has the form:

$$\psi_4(s) = \frac{12 \ln s}{5s} - \frac{4}{5s^3} + \frac{18}{5s^2} - \frac{13}{5s} - \frac{3}{10}. \tag{8}$$

The variance of this projection under H_0 is equal to

$$\Delta_4^2 = \int_1^\infty \psi_4^2(s) \frac{1}{s^2} ds = \frac{271}{52500} \approx 0.005.$$

Therefore the kernel Ψ_4 is centered, non-degenerate and bounded. Due to Hoeffding’s theorem on asymptotic normality of U -statistics (Hoeffding 1948; Korolyuk and Borovskikh 1994), we have that:

Theorem 4 *Under the null hypothesis as $n \rightarrow \infty$ the statistic $\sqrt{n}I_n^{(4)}$ is asymptotically normal so that*

$$\sqrt{n}I_n^{(4)} \xrightarrow{d} \mathcal{N}\left(0, \frac{271}{2100}\right).$$

The large deviation asymptotics of the sequence of statistics $I_n^{(4)}$ under H_0 follows from the following result. It was derived using the theorem on large deviations (see again Nikitin and Ponikarov 1999; DasGupta 2008; Nikitin 2010), applied to the centered, bounded and non-degenerate kernel Ψ_4 .

Theorem 5 *For $a > 0$*

$$\lim_{n \rightarrow \infty} n^{-1} \ln P(I_n^{(4)} > a) = -f_I^{(4)}(a),$$

where the function $f_I^{(4)}$ is continuous for the sufficiently small $a > 0$, and

$$f_I^{(4)}(a) \sim \frac{a^2}{50\Delta_4^2} = 3.875 a^2, \quad \text{as } a \rightarrow 0.$$

2.5 The local Bahadur efficiency of $I_n^{(4)}$

For this case the limit in probability under H_1 has the following asymptotics

$$b_1(\theta) \sim 5\theta \int_1^\infty \psi_4(s)h_1(s)ds,$$

where again $h_1(s) = \frac{\partial}{\partial \theta} g_1(s, \theta) |_{\theta=0}$ and $\psi_4(s)$ is the projection from (8). Therefore for the first Ley–Paindaveine alternative we have

$$b_1(\theta) \sim 5\theta \int_1^\infty \left(\frac{9}{8s^2} + \frac{3 \ln s}{2s} - \frac{1}{s} - \frac{1}{4} \right) \left(\frac{s-2}{s^3} \right) \frac{ds}{s^2} \sim \frac{\theta}{12}, \quad \theta \rightarrow 0.$$

and the local exact slope of the sequence $I_n^{(4)}$ as $\theta \rightarrow 0$ admits the representation

$$c_1(\theta) = b_1^2(\theta)/(25\Delta_4^2) \sim \frac{5}{66} \theta^2, \theta \rightarrow 0.$$

Table 2 Local Bahadur efficiencies for $I_n^{(4)}$

Alternative	Efficiency
Ley–Paindaveine 1	0.930
Ley–Paindaveine 2	0.961
log-Weibull	0.746

Table 3 Comparative table of local efficiencies for the statistic $I_n^{(k)}$

Alternative	Efficiency		
	$k = 3$	$k = 4$	\max_k
Ley–Paindaveine 1	0.909	0.930	0.930 for $k = 4$
Ley–Paindaveine 2	0.905	0.961	0.961 for $k = 4$
log-Weibull	0.787	0.746	0.821 for $k = 2$

The Kullback-Leibler “distance” for this alternative was already found above, and it satisfies $K_1(\theta) \sim \theta^2/24, \theta \rightarrow 0$. Thus the local efficiency of the test is

$$e_1^B(I) = \lim_{\theta \rightarrow 0} \frac{c_1(\theta)}{2K_1(\theta)} \approx 0.930.$$

For other alternatives the calculations are similar. Omitting the details, let us gather the values of the local BE for this case in the Table 2.

In Table 3 we present the efficiencies from Tables 1 and 2 gathered with maximal values of efficiencies against presumed alternatives.

3 Kolmogorov-type statistic $D_n^{(k)}$

Now we consider the Kolmogorov type statistic (3). For a fixed t the difference $H_n(t) - F_n(t)$ is a family of U -statistics with the kernels, depending on $t \geq 1$:

$$\mathcal{E}_k(X_{i_1}, \dots, X_{i_k}; t) = \mathbf{1}(X_{(k, \{i_1, \dots, i_k\})} / X_{(k-1, \{i_1, \dots, i_k\})} < t) - \frac{1}{k} \sum_{l=1}^k \mathbf{1}(X_l < t).$$

The projection of this kernels $\xi_k(s; t)$ for a fixed $t \geq 1$ has the form:

$$\begin{aligned} \xi_k(s; t) &:= E(\mathcal{E}_k(X_1, \dots, X_k) \mid X_k = s) \\ &= \mathbb{P}(X_{(k, \{1, \dots, k-1, s\})} / X_{(k-1, \{1, \dots, k-1, s\})} < t) - \frac{1}{k} \mathbf{1}\{s < t\} - \frac{k-1}{k} \mathbb{P}\{X_1 < t\}. \end{aligned}$$

It remains to calculate the first term. For this purpose like in the previous cases, we write the decomposition

$$\mathbb{P}(X_{k, \{1, \dots, k-1, s\}} / X_{k-1, \{1, \dots, k-1, s\}} < t) = \mathbb{P}_1 + \mathbb{P}_2 + \mathbb{P}_3,$$

where $\mathbb{P}_i, i = 1, 2, 3$, are the initial probabilities, computed in one of the following cases:

- (1) Let the sample units take places as follows: $X_1 < X_2 < \dots < X_{k-1} < s$. Then the probability expresses as

$$\begin{aligned} \mathbb{P}_1 &= (k - 1)! \mathbb{P} \left(\frac{s}{X_{k-1}} < t, X_1 < X_2 < \dots < X_{k-1} < s \right) \\ &= (k - 1)! \mathbf{1}(s \geq t) \mathbb{P} \left(\frac{s}{t} < X_{k-1} < s, X_1 < X_2 < \dots < X_{k-1} \right) \\ &\quad + (k - 1)! \mathbf{1}(s < t) \mathbb{P}(X_1 < X_2 < \dots < X_{k-1} < s) \\ &= \mathbf{1}(s \geq t) \left(F^{k-1}(s) - F^{k-1} \left(\frac{s}{t} \right) \right). \end{aligned}$$

- (2) The sample units are $X_1 < X_2 < \dots < X_{k-2} < s < X_{k-1}$, then for this case we have:

$$\begin{aligned} \mathbb{P}_2 &= (k - 1)! \mathbb{P} \left(\frac{X_{k-1}}{s} < t, X_1 < X_2 < \dots < X_{k-2} < s < X_{k-1} \right) \\ &= (k - 1)! \mathbb{P}(s < X_{k-1} < st, X_1 < X_2 < \dots < X_{k-2} < s) \\ &= (k - 1)! \frac{F^{k-2}(s)}{(k - 2)!} (F(st) - F(s)) = \frac{(k - 1)}{s} \left(1 - \frac{1}{s} \right)^{k-2} \left(1 - \frac{1}{t} \right). \end{aligned}$$

- (3) In the last case let s be situated on l -th place ($1 \leq l \leq k - 2$) in the variational series of the sample X_1, \dots, X_{k-2} . Then the required probability transforms into:

$$\begin{aligned} \mathbb{P}_3 &= (k - 1)! \mathbb{P} \left(\frac{X_{k-1}}{X_{k-2}} < t, X_1 < \dots < s < \dots < X_{k-2} < X_{k-1} \right) \\ &= \left(1 - \frac{1}{t} \right) C_{k-1}^{l-1} (1 - F(s))^{k-j} F^{j-1}(s), \quad 1 \leq l \leq k - 2. \end{aligned}$$

Combining these results we get that the first term in the projection is equal to:

$$\begin{aligned} &\mathbb{P}(X_{(k, \{1, \dots, k-1, s\})} / X_{(k-1, \{1, \dots, k-1, s\})} < t) \\ &= \mathbf{1}(s \geq t) \left(F^{k-1}(s) - F^{k-1} \left(\frac{s}{t} \right) \right) + \left(1 - \frac{1}{t} \right) \sum_{l=1}^{k-1} C_{k-1}^{l-1} (1 - F(s))^{k-j} F^{j-1}(s). \end{aligned}$$

Again we can see that the last sum can be simplified as

$$\sum_{l=1}^{k-1} C_{k-1}^{l-1} (1 - F(s))^{k-j} F^{j-1}(s) = 1 - F^{k-1}(s).$$

Thus the initial probability is equal to

$$\mathbb{P}(X_{(k, \{1, \dots, k-1, s\})} / X_{(k-1, \{1, \dots, k-1, s\})} < t) = \frac{1}{t} (F^{k-1}(s) - 1) - \mathbf{1}(s \geq t) F^{k-1} \left(\frac{s}{t} \right).$$

Hence we get the final expression for the projection of the family of kernels $\mathcal{E}_k(\cdot, t)$:

$$\xi_k(s; t) = \frac{1}{t} \left(\left(1 - \frac{1}{s} \right)^{k-1} - \frac{1}{k} \right) - \mathbf{1}(s \geq t) \left(\left(1 - \frac{t}{s} \right)^{k-1} - \frac{1}{k} \right). \tag{9}$$

It is easy to show that $E(\xi_k(X; t)) = 0$. After some calculations we get that the variance of this projection under H_0 is for any t

$$\begin{aligned} \delta_k^2(t) &= \frac{t+1}{(2k-1)t^2} + \frac{t-1}{k^2t^2} - \sum_{i=0}^{k-1} \frac{(-1)^j 2(k-1)!(k-1)!}{(k+j)!(k-j-1)!} t^{j-1} \\ &+ (-1)^{k+1} \frac{2(k-1)!(k-1)!}{(2k-1)!} t^{k-2} F^{2k-1}(t) - \frac{2}{k^2t} F^k(t). \end{aligned}$$

3.1 Kolmogorov-type statistic $D_n^{(3)}$

In the case $k = 3$ the projection of the family of kernels $\mathcal{E}_3(X, Y, Z; t)$, namely $\xi_3(s; t) := E(\mathcal{E}_3(X, Y, Z; t) \mid X = s)$ is equal to:

$$\xi_3(s; t) = \frac{1}{t} \left(\frac{1}{s^2} - \frac{2}{s} + \frac{2}{3} \right) - \mathbf{1}\{s \geq t\} \left(\frac{t^2}{s^2} - \frac{2t}{s} + \frac{2}{3} \right). \tag{10}$$

Now we calculate variances of these projections $\delta_3^2(t)$ under H_0 . Elementary calculations show that

$$\delta_3^2(t) = \frac{1}{45t^4} (4t^3 + 4t^2 - 15t + 7).$$

Hence our family of kernels $\mathcal{E}_3(X, Y, Z; t)$ is non-degenerate in the sense of Nikitin (2010) and

$$\delta_3^2 = \sup_{t \geq 1} \delta_3^2(t) = 0.035.$$

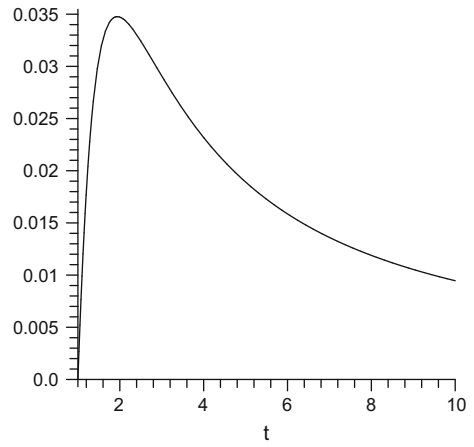
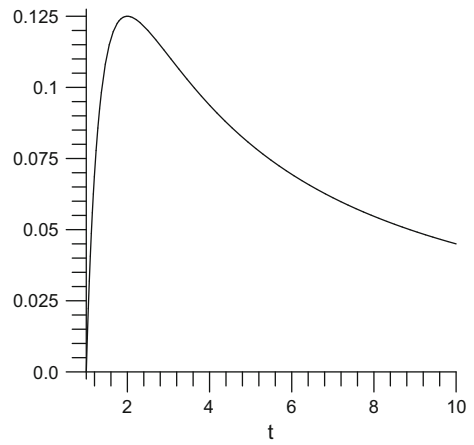
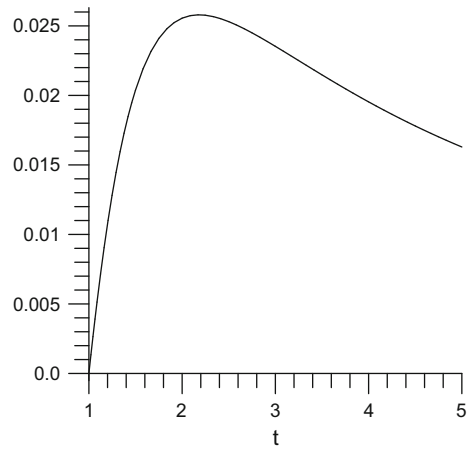
This value will be important in the sequel when calculating the large deviation asymptotics (Figs. 1, 2, 3).

The limiting distribution of the statistic $D_n^{(3)}$ is unknown. Using methods of Silverman (1983), one can show that the U -empirical process

$$\eta_n(t) = \sqrt{n} (H_n(t) - F_n(t)), \quad t \geq 1,$$

weakly converges in $D(1, \infty)$ as $n \rightarrow \infty$ to certain centered Gaussian process $\eta(t)$ with calculable covariance. Then the sequence of statistics $\sqrt{n}D_n^{(3)}$ converges in distribution to the rv $\sup_{t \geq 1} |\eta(t)|$ but currently it is impossible to find explicitly its distribution. Hence it is reasonable to determine the critical values for statistics $D_n^{(3)}$ by simulation.

Now we obtain the logarithmic large deviation asymptotics of the sequence of statistics $D_n^{(3)}$ under H_0 . The family of kernels $\{\mathcal{E}_3(X, Y, Z; t), t \geq 0\}$ is not only

Fig. 1 Plot of the function $\delta_{\frac{2}{3}}^2(t)$ **Fig. 2** Plot of the function $b_1(t, \theta)$, Ley–Paindaveine 1 alternative**Fig. 3** Plot of the function $\delta_{\frac{2}{4}}^2(t)$ 

centered but bounded. Using results from Nikitin (2010) on large deviations for the supremum of non-degenerate U -statistics, we obtain the following result.

Theorem 6 For $a > 0$

$$\lim_{n \rightarrow \infty} n^{-1} \ln P(D_n^{(3)} > a) = -f_D^{(3)}(a),$$

where the function $f_D^{(3)}$ is continuous for the sufficiently small $a > 0$, moreover

$$f_D^{(3)}(a) = (18\delta_3^2)^{-1} a^2 (1 + o(1)) \sim 1.598 a^2, \text{ as } a \rightarrow 0.$$

3.2 The local Bahadur efficiency of $D_n^{(3)}$

To evaluate the efficiency, first consider again the first Ley–Paindaveine alternative with the d.f. $G_1(x, \theta)$, $\theta \geq 0$, $x \geq 1$ given above. By the Glivenko–Cantelli theorem for U -statistics (Janssen 1988) the limit in probability under the alternative for statistics $D_n^{(3)}$ is equal to

$$b_1(\theta) := \sup_{t \geq 1} |b_1(t, \theta)| = \sup_{t \geq 1} |P_\theta(X_{(3,3)}/X_{(2,3)} < t) - G(t, \theta)|.$$

It is not difficult to show that

$$b_1(t, \theta) \sim 3\theta \int_1^\infty \xi_3(s; t) h_1(s) ds,$$

where again $h_1(s) = \frac{\partial}{\partial \theta} g_1(s, \theta) |_{\theta=0}$ and $\xi_3(s; t)$ is the projection defined above in (10). Hence for the first Ley–Paindaveine alternative we have for $t \geq 1$:

$$b_1(t, \theta) \sim \frac{t - 1}{2t^2} \theta, \quad \theta \rightarrow 0.$$

Thus $b_1(\theta) = \sup_{t \geq 1} |b_1(t, \theta)| \sim 0.125 \theta$, and it follows that the local exact slope of the sequence of statistics D_n admits the representation:

$$c_1(\theta) \sim b_1^2(\theta)/(9\delta_3^2) \sim 0.05 \theta^2, \quad \theta \rightarrow 0.$$

The Kullback–Leibler information in this case is given by (7). Hence the local Bahadur efficiency of our test is $e_1^B(D) = 0.599$.

Next we take the second Ley–Paindaveine distribution, where calculations are similar, and the local BE is equal to 0.689. In the case of the log–Weibull density we find that the local BE is 0.467.

We collect the values of local BE in the Table 4.

Table 4 Local Bahadur efficiencies for $D_n^{(3)}$

Alternative	Efficiency
Ley–Paindaveine 1	0.599
Ley–Paindaveine 2	0.689
log-Weibull	0.467

3.3 Kolmogorov-type statistic $D_n^{(4)}$

In the case $k = 4$ the projection of the family of kernels $\mathcal{E}_4(X, Y, Z, W; t)$, is equal to:

$$\xi_4(s; t) = \frac{1}{t} \left(\left(1 - \frac{1}{s}\right)^3 - \frac{1}{4} \right) - \mathbf{1}\{s \geq t\} \left(-\left(\frac{t}{s}\right)^3 + 3\left(\frac{t}{s}\right)^2 - \frac{3t}{s} + \frac{3}{4} \right).$$

Therefore we get that variances of these projections $\delta_4^2(t)$ under H_0

$$\delta_4^2(t) = \frac{1}{560t^5} (45t^4 + 45t^3 - 252t^2 + 224t - 62).$$

Hence our family of kernels $\mathcal{E}_4(X, Y, Z, W; t)$ is non-degenerate in the sense of Nikitin (2010) and

$$\delta_4^2 = \sup_{t \geq 1} \delta_4^2(t) = 0.026.$$

The limiting distribution of the statistic $D_n^{(4)}$ is unknown as in the previous section.

The logarithmic large deviation asymptotics of the sequence of statistics $D_n^{(4)}$ under H_0 is showed in the next theorem.

Theorem 7 For $a > 0$

$$\lim_{n \rightarrow \infty} n^{-1} \ln P(D_n^{(4)} > a) = -f_D^{(4)}(a),$$

where the function $f_D^{(4)}$ is continuous for the sufficiently small $a > 0$, moreover

$$f_D^{(4)}(a) = (32\delta_4^2)^{-1} a^2 (1 + o(1)) \sim 1.211 a^2, \text{ as } a \rightarrow 0.$$

3.4 The local Bahadur efficiency of $D_n^{(4)}$

In the Table 5 we collect the calculated efficiencies for the statistic $D_n^{(k)}$ joined with results from the Table 4 and with the maximal values of efficiencies against our alternatives.

We observe that the efficiencies for the Kolmogorov-type test are lower than for the integral test. However, it is the usual situation when testing goodness-of-fit (Nikitin 1995; Rank 1999; Nikitin 2010).

Table 5 Comparative table of local efficiencies for statistic $D_n^{(k)}$

Alternative	Efficiency		
	$k = 3$	$k = 4$	\max_k
Ley–Paindaveine 1	0.599	0.654	0.674 for $k = 6$
Ley–Paindaveine 2	0.689	0.767	0.790 for $k = 5$
log-Weibull	0.467	0.472	0.472 for $k = 4$

Table 6 Critical values for the statistic $D_n^{(3)}$

$n \setminus \alpha$	0.1	0.05	0.01
10	0.39	0.42	0.56
20	0.30	0.32	0.36
30	0.25	0.27	0.31
40	0.23	0.25	0.28
50	0.21	0.23	0.26
100	0.17	0.18	0.21

Table 7 Critical values for the statistic $D_n^{(4)}$

$n \setminus \alpha$	0.1	0.05	0.01
10	0.47	0.50	0.54
20	0.37	0.40	0.44
30	0.33	0.35	0.38
40	0.30	0.33	0.35
50	0.28	0.30	0.33
100	0.24	0.26	0.29

3.5 Critical values

Tables 6 and 7 shows the critical values of the null distribution of $D_n^{(3)}$ and $D_n^{(4)}$ for significance levels $\alpha = 0.1, 0.05, 0.01$ and specific sample sizes n . Each entry is obtained by using the Monte-Carlo simulation methods with 10,000 replications.

4 Power comparison

We recall computation formulae for statistics $I_n^{(k)}$ and $D_n^{(k)}$ for $k = \{3, 4\}$:

$$\begin{aligned}
 I_n^{(k)} &= \int_1^\infty (H_n(t) - F_n(t)) dF_n(t) \\
 &= \binom{n}{k+1}^{-1} \sum_{1 \leq i_1 < \dots < i_{k+1} \leq n} \frac{1}{k+1} \sum_{\pi(i_1, \dots, i_{k+1})} \mathbf{1}(X_{(k, \{i_1, \dots, i_k\})} / X_{(k-1, \{i_1, \dots, i_k\})} < X_{i_{k+1}}) - \frac{1}{2},
 \end{aligned}$$

where $\pi(i_1, \dots, i_{k+1})$ means all permutations of different indices from $\{i_1, \dots, i_{k+1}\}$.

$$D_n^{(k)} = \sup_{t \geq 1} |H_n(t) - F_n(t)|$$

$$= \sup_{t \geq 1} \left| \binom{n}{k}^{-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} \left[\mathbf{1}(X_{(k, \{i_1, \dots, i_k\})} / X_{(k-1, \{i_1, \dots, i_k\})} < t) - \frac{1}{k} \sum_{l=1}^k \mathbf{1}(X_l < t) \right] \right|.$$

This section presents results of a Monte-Carlo study to compare powers of new tests with the widely applied for these types of hypotheses Kolmogorov-Smirnov (KS) and Cramer-von Mises (CvM) tests. The comparison is done for the size $n = 20$ and for the significance level $\alpha = 0.05$. All calculation were done using JAVA (The Apache Commons Mathematics Library) and R package with 10,000 replications. We consider following distributions as alternatives for the Pareto distribution:

- 1) the Gamma alternative $\Gamma(\theta)$ with the density $(\Gamma(\theta))^{-1} x^{\theta-1} \exp(-x)$;
- 2) the log-normal law $LN(\theta)$ with the density $(\theta x)^{-1} (2\pi)^{-1/2} \exp(-(\log x)^2 / 2\theta^2)$;
- 3) the first Ley–Paindaveine alternative $LP1(\theta)$ with the d.f. $(1 - \frac{1}{x}) \exp(-\theta/x)$, $\theta \geq 0, x \geq 1$;
- 4) log-Weibull alternative with the d.f. $1 - e^{-(\ln x)^{\theta+1}}$, $\theta \in (0, 1), x \geq 1$;
- 5) the Weibull distribution $W(\theta)$ with the density $\theta x^{\theta-1} \exp(-x^\theta)$.

The KS and CvM tests are not applicable to composite hypothesis, so first we estimate parameter λ with its maximum likelihood estimator (MLE) $\hat{\lambda} = n(\sum_{k=1}^n \ln X_k)^{-1}$, then calculate the critical values of corrected test and powers for the sample $n = 20$ using the Monte-Carlo procedure. Powers are given in the Table 8.

We observe that powers of our tests correspond to local Bahadur efficiencies for considerable alternatives. In whole we can notice that our new statistics in comparison with classical tests more favorable to alternatives with the density shapes similar to the Pareto distribution, for example like the first Ley–Paindaveine alternative. However they are less responsive to close alternatives but with another shapes (for example when density have some twists differ to the Pareto distribution), for example gamma and log-Weibull alternatives.

Table 8 Comparative table of the power simulation for different statistics

Statistics	$I_n^{(3)}$	$I_n^{(4)}$	$D_n^{(3)}$	$D_n^{(4)}$	KS	CvM
$\Gamma(2)$	0.7311	0.8505	0.4291	0.2277	0.8602	0.9605
$LN(1)$	0.7090	0.7126	0.7192	0.7390	0.5573	0.6391
$LN(1.5)$	0.6984	0.6997	0.7305	0.7512	0.5685	0.6540
$LP1(0.25)$	0.4618	0.4722	0.4894	0.5001	0.0712	0.2471
$LP1(0.5)$	0.6816	0.7013	0.6899	0.6975	0.0887	0.2903
$LW(0.25)$	0.2551	0.2429	0.0835	0.0178	0.0213	0.2130
$LW(0.5)$	0.5816	0.5677	0.2526	0.0622	0.0442	0.2736
$W(2)$	1	1	0.9992	0.9950	0.9904	1

Table 9 Goodness-of-fit tests for fitted Pareto models based on MLE $\lambda = 0.764$ and $\sigma = 1.5$ for statistics $I_n^{(k)}, D_n^{(k)}$

Statistic	$I_n^{(3)}$	$I_n^{(4)}$	$D_n^{(3)}$	$D_n^{(4)}$
p value	0.73	0.44	0.32	0.27

5 Application to the real data

In this section we apply our tests to the real data example from [Hogg and Klugman \(1984\)](#), where they are discussed in detail. The data set represents the losses due to wind-related catastrophes, 1977, rounded to the nearest million dollars and involved more than \$2 million:

2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 4 4 4 5
 5 5 5 6 6 6 6 8 8 9 15 17 22 23 24 24 25 27 32 43

These data widely analyze in the literature, see [Brazauskas and Serfling \(2003\)](#) for detail, the new goodness-of-fit tests were proposed in [Rizzo \(2009\)](#). First we apply the same to [Brazauskas and Serfling \(2003\)](#) and [Rizzo \(2009\)](#) the data de-grouping method. The necessity of this method appear as a consequence of the initial data rounding and make from discrete grouping the uniform distributed data. Put

$$X_k = \left(1 - \frac{k}{m + 1}\right) A + \frac{k}{m + 1} B, k = 1, \dots, m,$$

where (A, B) is a grouping interval with m observations. According to [Brazauskas and Serfling \(2003\)](#) for example to observations corresponds to 2 we conditionally take as (A, B) the interval $(1.5, 2.5)$.

We tested the null-hypothesis $H_0 : X$ has the Pareto distribution with the scale parameter $\sigma = 1.5$ and MLE of the tail parameter $\lambda = 0.764$. Such special parameters consider in [Brazauskas and Serfling \(2003\)](#), [Hogg and Klugman \(1984\)](#) and in [Rizzo \(2009\)](#), [Philbrick and Jurschak \(1981\)](#) applied $\sigma = 2.0$.

Applying our tests to these data, we get in the [Table 9](#) the following p-values of test statistics $I_n^{(k)}$ and $D_n^{(k)}$, based on 10,000 simulations.

So we conclude that our tests do not reject the null-hypothesis. It correspond to results of [Brazauskas and Serfling \(2003\)](#), [Rizzo \(2009\)](#).

6 Conditions of the local asymptotic optimality

In this section we are interested in conditions of the local asymptotic optimality (LAO) in Bahadur sense for both sequences of statistics $I_n^{(k)}$ and $D_n^{(k)}$. This means to describe the local structure of alternatives for which the given statistic has maximal potential local efficiency so that the relation

$$c_T(\theta) \sim 2K(\theta), \quad \theta \rightarrow 0,$$

holds ([Nikitin 1995](#); [Nikitin and Tchirina 1996](#)). Such alternatives form the domain of LAO for the given sequence of statistics.

Consider functions

$$H(x) = G'_\theta(x, \theta) |_{\theta=0}, \quad h(x) = g'_\theta(x, \theta) |_{\theta=0} .$$

We will assume that the following regularity conditions are true (see also Nikitin and Tchirina 1996):

$$\int_1^\infty h^2(x)x \, dx < \infty \quad \text{where} \quad h(x) = H'(x), \tag{11}$$

$$\frac{\partial}{\partial \theta} \int_1^\infty g(x, \theta) \ln x \, dx |_{\theta=0} = \int_1^\infty h(x) \ln x \, dx. \tag{12}$$

Denote by \mathcal{G} the class of densities $g(x, \theta)$ with d.f.'s $G(x, \theta)$, satisfying the regularity conditions (11)–(12). We are going to deduce the LAO conditions in terms of the function $h(x)$.

Recall that for alternative densities from \mathcal{G} the following asymptotics is valid:

$$2K(\theta) \sim \theta^2 \left[\int_1^\infty h^2(x)x \, dx - \left(\int_1^\infty h(x) \ln x \, dx \right)^2 \right], \quad \theta \rightarrow 0.$$

6.1 LAO conditions for $I_n^{(k)}$

First consider the integral statistic $I_n^{(k)}$ with the kernel $\Psi_k(X_1, \dots, X_{k+1})$ and its projection $\psi_k(x)$ from (4). Let introduce the auxiliary function

$$h_0(x) = h(x) - \frac{(\ln x - 1)}{x^2} \int_1^\infty h(u) \ln u \, du.$$

Simple calculations show that

$$\begin{aligned} \int_1^\infty h^2(x)x^2 \, dx - \left(\int_1^\infty h(x) \ln x \, dx \right)^2 &= \int_1^\infty h_0^2(x)x^2 \, dx, \\ \int_1^\infty \psi_k(x)h(x) \, dx &= \int_1^\infty \psi_k(x)h_0(x) \, dx. \end{aligned}$$

Hence the local asymptotic efficiency takes the form

$$\begin{aligned} e^B(I_n^{(k)}) &= \lim_{\theta \rightarrow 0} b_I^2(\theta) / \left((k + 1)^2 \Delta_k^2 \cdot 2K(\theta) \right) \\ &= \left(\int_1^\infty \psi_k(x)h_0(x) \, dx \right)^2 / \left(\int_1^\infty \psi_k^2(x) \frac{dx}{x^2} \cdot \int_1^\infty h_0^2(x)x^2 \, dx \right). \end{aligned}$$

Table 10 Examples of LAO alternative density $g(x, \theta)$ for statistic $I_n^{(k)}$

Alternative density $g(x, \theta)$ as $\theta \rightarrow +0, x \geq 1$	
$k = 3$	$g(x, \theta) = \frac{1}{x^2} \left(1 + \theta \left(\frac{9}{8x^2} + \frac{3 \ln x}{2x} - \frac{1}{x} - \frac{1}{4} \right) \right)$
$k = 4$	$g(x, \theta) = \frac{1}{x^2} \left(1 + \theta \left(\frac{12 \ln x}{5x} - \frac{4}{5x^3} + \frac{18}{5x^2} - \frac{13}{5x} - \frac{3}{10} \right) \right)$

By Cauchy-Schwarz inequality we obtain that the expression in the right-hand side is equal to 1 iff $h_0(x) = C_1 \psi_k(x) \frac{1}{x^2}$ for some constant $C_1 > 0$, so that

$$h(x) = (C_1 \psi_k(x) + C_2(\ln x - 1)) \frac{1}{x^2} \quad \text{for some constants } C_1 > 0 \text{ and } C_2. \quad (13)$$

The set of distributions for which the function $h(x)$ has such form generate the domain of LAO in the class \mathcal{G} . The simplest examples of such alternatives density $g(x, \theta)$ for small $\theta > 0$ is given by the Table 10.

6.2 LAO conditions for $D_n^{(k)}$

Now let consider the Kolmogorov type statistic $D_n^{(k)}$ with the family of kernels Ξ_k and their projections $\xi_k(x; t)$ from (9). After simple calculations we get

$$\int_1^\infty \xi_k(x; t) h(x) dx = \int_1^\infty \xi_k(x; t) h_0(x) dx, \quad \forall t \in [1, \infty).$$

Hence the local efficiency takes the form

$$\begin{aligned} e^B(D_n^{(k)}) &= \lim_{\theta \rightarrow 0} \left[\frac{b_D^2(\theta) / \sup_{t \geq 1} (k^2 \delta_k^2(t)) \cdot 2K(\theta)}{\sup_{t \geq 1} \left(\int_1^\infty \xi_k(x; t) h_0(x) dx \right)^2} \right] \\ &= \frac{\sup_{t \geq 1} \left(\int_1^\infty \xi_k(x; t) h_0(x) dx \right)^2}{\sup_{t \geq 1} \left(\int_1^\infty \xi_k^2(x; t) \frac{dx}{x^2} \cdot \int_1^\infty h_0^2(x) x^2 dx \right)} \leq 1. \end{aligned}$$

We can apply once again the Cauchy-Schwarz inequality to the numerator in the last ratio. It follows that the sequence of statistics D_n is the locally asymptotically optimal, and $e^B(D_n^{(k)}) = 1$ iff

$$h(x) = (C_3 \xi_k(x; t_0) + C_4(\ln x - 1)) \cdot \frac{1}{x^2} \quad \text{for } t_0 = \arg \sup_{t \geq 1} \delta_k^2(t)$$

and some constants $C_3 > 0$ and C_4 .

Distributions with such $h(x)$ form the domain of LAO in the class \mathcal{G} . The simplest examples are given in the Table 11.

Table 11 Examples of LAO alternative density $g(x, \theta)$ for statistic $D_n^{(k)}$

Alternative densities $g(x, \theta)$ as $\theta \rightarrow +0$, $x \geq 1$	
$k = 3$	$g(x, \theta) = \frac{1}{x^2} \left(1 + \theta \left(\frac{1}{t_1} \left(\frac{1}{x^2} - \frac{2}{x} + \frac{2}{3} \right) - \mathbf{1}\{x \geq t_1\} \left(\left(\frac{t_1}{x} \right)^2 - \frac{2t_1}{x} + \frac{2}{3} \right) \right) \right)$ $t_1 = \arg \max_{t \geq 1} \left(\frac{1}{45t^4} (4t^3 + 4t^2 - 15t + 7) \right) \approx 1.9395$
$k = 4$	$g(x, \theta) = \frac{1}{x^2} \left(1 + \theta \left(\frac{1}{t_2} \left(\left(1 - \frac{1}{x} \right)^3 - \frac{1}{4} \right) - \mathbf{1}\{x \geq t_2\} \left(- \left(\frac{t_2}{x} \right)^3 + 3 \left(\frac{t_2}{x} \right)^2 - \frac{3t_2}{x} + \frac{3}{4} \right) \right) \right)$ $t_2 = \arg \max_{t \geq 1} \left(\frac{1}{560t^5} (45t^4 + 45t^3 - 252t^2 + 224t - 62) \right) \approx 2.1810$

7 Conclusion

We constructed two new tests for goodness-of-fit testing for the Pareto distribution based on the new characterization for the Pareto distribution. We describe their limit distribution and large deviations. The Bahadur efficiency for some alternatives has been obtained and it turned out reasonably high. Also we derived the conditions of local optimality for our tests. These tests were compared with some commonly used goodness-of-fit tests and it can be noted that in most cases our tests are more powerful. They can be of some use in statistical research, especially when the alternative is close to the alternative from the LAO class.

Acknowledgments The authors express their deep gratitude to the Referees and the Associate Editor for their useful suggestions for the improvement of the paper.

References

- Ahsanullah M (1977) A characteristic property of the exponential distribution. *Ann Stat* 5(3):580–582
- Ahsanullah M (1989) On characterizations of the uniform distribution based on functions of order statistics. *Allgarh J Stat* 9:1–6
- Arnold BC (1983) Pareto distributions. International Co-operative Publishing House, Fairland, MD
- Bahadur RR (1971) Some limit theorems in statistics. SIAM, Philadelphia
- Beirlant J, de Wet T, Goegebeur Y (2006) A goodness-of-fit statistic for the Pareto-type behavior. *J Comp Appl Math* 186:99–116
- Brazauskas V, Serfling R (2003) Favorable estimators for fitting Pareto models: a study using goodness-of-fit measures with actual data. *ASTIN Bull* 33(2):365–381
- DasGupta A (2008) Asymptotic theory of statistics and probability. Springer, New York
- Gulati S, Shapiro S (2008) Goodness of fit tests for the Pareto distribution. In: Vonta F, Nikulin M, Limnios N, Huber C (eds) Statistical models and methods for biomedical and technical systems. Birkhäuser, Boston, pp 263–277
- Helmers R, Janssen P, Serfling R (1988) Glivenko–Cantelli properties of some generalized empirical DF's and strong convergence of generalized L-statistics. *Prob Theory Relat Fields* 79:75–93
- Hoeffding W (1948) A class of statistics with asymptotically normal distribution. *Ann Math Stat* 19:293–325
- Hogg RV, Klugman SA (1984) Loss distributions. Wiley, New York
- Janssen PL (1988) Generalized empirical distribution functions with statistical applications. Limburgs Universitair Centrum, Diepenbeek
- Jovanovic M, Milosevic B, Obradovic M (2014) Goodness of fit tests for Pareto distribution based on a characterization and their asymptotics. [arXiv:1310.5510](https://arxiv.org/abs/1310.5510). Accepted for publication in *Statistics* doi: 10.1080/02331888.2014.919297

- Kleiber C, Kotz SA (2003) Statistical size distributions in economics and actuarial sciences. Wiley
- Korolyuk VS, Borovskikh YV (1994) Theory of U -statistics. Kluwer, Dordrecht
- Ley C, Paindaveine D (2008) Le Cam optimal tests for symmetry against Ferreira and Steel's general skewed distribution. *J Nonparam Stat* 21:943–967
- Litvinova VV (2004) Asymptotic properties of goodness-of-fit and symmetry tests based on characterizations. Dissertation, Saint-Petersburg University
- Martynov GV (2009) Cramér-von Mises test for the Weibull and Pareto Distributions. In: Proceedings of Dobrushin Intern Conf, Moscow, pp 117–122
- Nikitin Y (1995) Asymptotic efficiency of nonparametric tests. Cambridge University Press, New York
- Nikitin YY (1996) Bahadur efficiency of a test of exponentiality based on a loss of memory type functional equation. *J Nonparametr Stat* 6(1):13–26
- Nikitin YY (2010) Large deviations of U -empirical Kolmogorov–Smirnov tests, and their efficiency. *J Nonparametr Stat* 22:649–668
- Nikitin YY, Peaucelle I (2004) Efficiency and local optimality of distribution-free tests based on U - and V -statistics. *Metron* LXII:185–200
- Nikitin YY, Ponikarov EV (1999) Rough large deviation asymptotics of Chernoff type for von Mises functionals and U -statistics. In: Proceedings of StPetersburg Math Soc, vol 7, pp 124–167. Engl transl (2001) *AMS Transl*, vol 2(203), pp 107–146
- Nikitin YY, Tchirina AV (1996) Bahadur efficiency and local optimality of a test for the exponential distribution based on the Gini statistic. *Stat Methods Appl* 5:163–175
- Nikitin YY, Volkova KY (2012) Asymptotic efficiency of goodness-of-fit test for power distribution based on Puri–Rubin characterization. *Zapisk Nauchnykh Semin POMI* 408:115–130
- Nikitin YY, Volkova KY (2010) Asymptotic efficiency of exponentiality tests based on order statistics characterization. *Georgian Math J* 17:749–763
- Philbrick SW, Jurschak J (1981) Discussion of “Estimating casualty insurance loss amount distributions”. In: Proceedings of the Casualty Actuarial Society, LXVIII
- Rank RF (1999) Statistische Anpassungstests und Wahrscheinlichkeiten grosser Abweichungen. Dissertation, University of Hannover
- Rizzo ML (2009) New goodness-of-fit tests for Pareto distribution. *Astin Bull* 39(2):691–715
- Silverman BW (1983) Convergence of a class of empirical distribution functions of dependent random variables. *Ann Probab* 11:745–751
- Wieand HS (1976) A condition under which the Pitman and Bahadur approaches to efficiency coincide. *Ann Stat* 4:1003–1011