Записки научных
семинаров ПОМИ
Том 442, 2015 г.

V. Bagdonavičius, M. Nikulin, A. Zerbet

ON OUTLIERS DETECTION FOR LOCATION-SCALE AND SHAPE-SCALE FAMILIES

Abstract. The problem of multiple upper outliers detection in samples from location-scale and shape-scale families is considered. A new test statistic is proposed. The critical values of the new test statistic are tabulated by simulation. The power of the new test and other available tests are compared by simulation.

§1. Introduction

The problem of outlier detection in a sample of observations is an old problem. Various criteria have been proposed and their performances studied in statistical literature. Most of them have been concentrated on the normal distribution (see Dixon [6], Grubbs [8], Большев, Смирнов [4], Большев, Убайдулаева [3], Tietjen and Moore [21], Barnett and Lewis [2], Balakrishnan [1]. For non-normal case, the bulk of the literature pertains to the exponential and gamma distributions, see, for example, Likes [14], Kabe [10], Lewis and Fiellerm [13], Chikkagoudar and Kunchur [5], Zerbet and Nikulin [22], Lalitha and Kumar [12], Kumar [11], Ibragimov and Khalfina [9], Pagurova [18]. Ю. В. Линник [15] collaborated also much on the problem of detection of outliers and on elimination of crude observations.

Problems of outliers in samples from the Weibull or extreme-value distributions are recognized to be of considerable practical importance (see Mann [17]). These distributions are special cases of shape-scale and location-scale families of distributions, respectively. They have often been used as pragmatic models in survival analysis, both in the biological and in the physical sciences, as well as in many other applications. A survey of the outlier detection procedures in the case of Weibull or extreme-value distributions using a location-scale models was given in Fung and Paul [7]. See also Mann, Scheuer and Fertig [16], Rosner [20].

Key words and phrases: location-scale family, non-spurious power, outliers, power comparison, shape-scale family, slippage alternative, Weibull and extreme-value distributions.

The probability distribution of a random variable T belongs to the *shape-scale family* of distributions with the c.d.f.

$$F_T(t) = F_0((t/\theta)^\nu) \quad \text{for all } t \in \mathbf{R}, \tag{1}$$

where F_0 is a *completely specified* c.d.f., $F_0(t) = 0$ for $t < 0$, and $\theta > 0$, $\nu > 0$ are *unknown scale and shape parameters.*

In the case of the *Weibull, loglogistic and lognormal distributions*

$$F_0(t) = 1 - e^{-t}, \quad \frac{t}{1+t}, \quad \Phi(\ln t), \tag{2}$$

respectively, where $\Phi(x)$ is the c.d.f. of the *standard normal distribution.*

The distribution of the random variable $X = \ln T$ belongs to the family of *location-scale distributions* with the c.d.f.

$$F_X(x) = F_0(e^{\frac{x-\mu}{\sigma}}) = F((x-\mu)/\sigma), \quad x \in \mathbf{R}; \tag{3}$$

here $\mu = \ln\theta$, $\sigma = 1/\nu$.

In the case of the baseline function F_0 given in (2) the *extreme values, logistic and normal distributions* are obtained, respectively.

So any test constructed for *location-scale family* can be used in the case of corresponding shape-scale family. It suffices to use logarithms of the sample members.

§2. The null hypothesis and the slippage alternative

Let us consider the null hypothesis

H_0: *the sample $X_1, X_2, \ldots, X_n$ derives from the location-scale family* (3) *with a specified baseline c.d.f. F and the parameters μ, σ.*

Under H_0 the p-quantile of the distribution is $x_p = \mu + \sigma F^{-1}(p)$. It implies that the random variable X has tendency to take greater values if the parameters μ increases. Upper quantiles $F^{-1}(p)$ of the baseline distribution are positive so the upper order statistics of X also have tendency to take greater values if σ increases.

So we consider the following *slippage alternative* of existence of r upper outliers:

$H_{r,\delta,\Delta}$: *$n-r$ from X_i have the distribution* (3) *with the parameters μ and σ and r from them have distribution* (3) *with the parameters $(\mu+\delta, \sigma\Delta)$, where $\delta \geqslant 0$, $\Delta \geqslant 1$, $(\delta, \Delta) \neq (0, 1)$.*

This alternative includes cases when one or both of the parameters μ or σ increase. The most natural alternatives are $H_{r,\delta,1}$. For example under *the accelerated failure time alternative covariates* influence only scale parameter and not the shape parameter.

Suppose that instead of location-scale families we consider scale-shape families. So instead of the data $X_1, \ldots, X_n$ we consider the data $T_1, \ldots, T_n$, $X_i = \ln T_i$. Set

$$F(x) = F_0(e^x), \quad \mu = \ln\theta, \quad \sigma = 1/\nu.$$

The null hypothesis and the alternative can be formulated in equivalent form:

H_0 : *the sample* $T_1, T_2, \ldots, T_n$ *derives from the scale-shape family* (1) *with a specified baseline c.d.f.* F_0 *and the parameters* μ,σ.

Under H_0 the p-quantile of the distribution is $t_p = \theta\,[F_0^{-1}(p)]^{1/\nu}$. It implies that the random variable T has tendency to take greater values if the parameters θ increases. Upper quantiles $F_0^{-1}(p)$ of the baseline distribution are greater than 1 so the upper order statistics of T also have tendency to take greater values if ν decreases. The slippage alternative is the following:

$H_{r,k,\Delta}$: $n - r$*from* T_i *have the same distribution as under* H_0 *and* r *from them have the distribution from the same scale-shape family with the parameters* $(\theta k, \nu/\Delta)$, *where* $k \geqslant 1$, $\Delta \geqslant 1$, $(k, \Delta) \neq (1, 1)$.

The most natural are location alternatives $H_{r,k,1}$.

§3. Available tests

Let $X_{(1)} \leqslant X_{(2)} \leqslant \ldots \leqslant X_{(n)}$ be an ordered sample from a specified location-scale family with location parameter μ and scale parameter σ. Let us introduce the most popular tests for detecting r upper outliers in location-scale or shape-scale families.

Dixon's tests, Dixon [6]. Three test statistics are mostly used:

$$D_{1r} = \frac{X_{(n)} - X_{(n-r)}}{X_{(n)} - X_{(1)}}, \quad D_{2r} = \frac{X_{(n)} - X_{(n-r)}}{X_{(n)} - X_{(2)}}, \quad D_{3r} = \frac{X_{(n)} - X_{(n-r)}}{X_{(n)} - X_{(3)}}.$$

The hypothesis H_0 is rejected if the test statistic takes large values.

Grubbs test, Grubbs [8]. The test statistic is

$$G = S^2_{n-r+1,\ldots,n}/S^2;$$

here

$$S^2 = \sum_{i=1}^{n} (X_{(i)} - \bar{X})^2, \quad \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_{(i)},$$

$$S^2_{n-r+1,\dots,n} = \sum_{i=1}^{n-r} (X_{(i)} - \bar{X}_{n-r+1,\dots,n})^2, \quad \bar{X}_{n-r+1,\dots,n} = \frac{1}{n-r} \sum_{i=1}^{n-r} X_{(i)}.$$

The hypothesis H_0 is rejected if the test statistic takes small values.

Grubbs-2 test is based on the statistic

$$G_r = \frac{\sum\limits_{i=n-r+1}^{n} X_{(i)} - r\bar{X}}{s_X}, \quad s_X^2 = \frac{1}{n} \sum_{j=1}^{n} (X_j - \bar{X})^2.$$

The hypothesis H_0 is rejected if the test statistic takes large values.

Mann's test, Mann [17]. The test statistic is

$$L = \sum_{i=n-r-1}^{n-1} l_i \Big/ \sum_{i=1}^{n-1} l_i$$

where

$$l_i = (X_{(i+1)} - X_{(i)})/(\mathbf{E}(Z_{(i+1)}) - \mathbf{E}(Z_{(i)})),$$

and

$$Z_{(i)} = (X_{(i)} - \mu)/\sigma.$$

The hypothesis H_0 is rejected if the test statistic takes small values. This test was initially proposed for detecting upper outliers in Weibull samples taking logarithms of the initial data $T_1, \dots, T_n$. This approach can also be used to shape-scale families of distributions because for such families the distribution of the test statistic under the null hypothesis is parameter-free.

Kumar's test, Kumar [11]. The test statistic

$$K_r = \frac{X_{(n)} - X_{(n-r)}}{\sum\limits_{i=2}^{n} (X_{(i)} - X_{(1)})}$$

was proposed for detecting upper outliers in exponential samples with unknown origin. This approach can be used to shape-scale family of distributions because for such families the distribution of the test statistic under the null hypothesis is parameter-free.

Tietjen–Moore test, Tietjen and Moore [21]. Tietjen and Moore proposed the statistic E_r for detecting upper outliers in normal samples. The test statistic is

$$E_r = \frac{\sum\limits_{i=1}^{n-r} (X_{(i)} - \bar{X}_r)^2}{\sum\limits_{i=1}^{n} (X_{(i)} - \bar{X})^2},$$

where

$$\bar{X}_r = \frac{1}{n-r} \sum_{i=1}^{n-r} X_{(i)}.$$

The hypothesis H_0 is rejected if the test statistic takes small values. This test can be used for location-scale families because for such families the distribution of the test statistic under the null hypothesis is parameter-free.

§4. The new test

We begin by noting that the mean of X_i is $\mathbf{E}\, X_i = \mu + a\sigma$ for all i under H_0 and for $n-r$ from X_i under $H_{r,\delta,\Delta}$, whereas $\mathbf{E}\, X_i = \mu + \delta + a\sigma\Delta$ for r from X_i under $H_{r,\delta,\Delta}$; here $a = \int\limits_{-\infty}^{\infty} x\, dF(x)$.

Set $\widetilde{l}_{i,n} = M_{i,n}(X_{(i+1)} - X_{(i)})$, where

$$M_{i,n} = f(F^{-1}(i/(n+1))), \quad f(x) = F'(x),$$

For example, in the case of the extreme values, logistic and normal distribution

$$M_{in} = -\Big(1 - \frac{i}{n+1}\Big) \ln\Big(1 - \frac{i}{n+1}\Big); \quad \frac{i}{n+1}\Big(1 - \frac{i}{n+1}\Big); \quad \varphi\Big(\Phi^{-1}\Big(\frac{i}{n+1}\Big)\Big),$$

respectively, where $\varphi(x)$ is the p.d.f. of the standard normal distribution.

The idea of the coefficient $M_{i,n}$ choice is derived (see Mann *et al.* [16]) from the following limit theorem (see Pyke [19]): under general assumptions the random variables $(n(Z_{(i+1)} - Z_{(i)})$ and $n(Z_{(j+1)} - Z_{(j)})$ are asymptotically (as $n \to \infty$, $i/n \to p_1 \in (0,1)$, $j/n \to p_2 \in (0,1)$) independent and converge in distribution to the exponential law with the means $1/f(F^{-1}(p_1))$ and $1/f(F^{-1}(p_2))$, respectively; $Z_{(i)}$ are the order statistics from the parameter-free distribution F. The distribution of the random vector $(\widetilde{l}_{1,n}/\sigma, \ldots, \widetilde{l}_{n-1,n}/\sigma)^T$ does not depend on unknown parameters. The components of this vector are almost equally distributed.

We propose the following test statistic

$$B(r,n) = \frac{(n-r)\sum_{k=n-r+1}^{n} X_{(k)} - r\sum_{i=1}^{n-r} X_{(i)}}{n\sum_{j=1}^{n-r-1} \widetilde{l}_{(j)}}.$$

Note that for the data without outliers

$$X_{(i)} = \sigma Z_{(i)} + \mu,$$

and

$$\widetilde{l}_{(i)} = M_{i,n}(X_{(i+1)} - X_{(i)}) = \sigma M_{i,n}(Z_{(i+1)} - Z_{(i)}),$$

so the distribution of the statistic $B(r,n)$ is parameter free under H_0.

Under the alternative $H_{r,\delta,\Delta}$ the r random variables from X_i which have tendency to take greater values than other ones, give the effect that order statistics under this alternative have tendency to take greater values than under the null hypothesis. This effect concerns mostly the upper order statistics. So in average the first term of the numerator of the test statistic should increase more then the second term going from the hypothesis to alternative. The distribution of the denominator should not change much. So the hypothesis should be rejected if the test statistic takes large values.

If scale-shape families are considered then the data $T_1, \dots, T_n$ should be transformed to the data $X_1, \dots, X_n$ by the transform $X_i = \ln T_i$.

Under the null hypothesis the distribution of each test statistics is parameter-free, so investigating by simulation the power of the tests the values $\mu = 0, \sigma = 1$ (or, equivalently, $\theta = \nu = 1$) could be taken when simulating the data under the null hypothesis, and different values of δ (or, equivalently, k) could be considered for fixed $\Delta = 1$ for one group of simple alternatives or different values of Δ could be considered for fixed $\delta = 0$ (or, equivalently, $k = 1$) for other group of simple alternatives. Various values of pairs (δ, Δ) (or (k, Δ)) could also be considered.

§5. Critical values of $B(r,n)$

In this section, for several shape-scale families, we give the critical values of the statistic $B(r,n)$ computed by simulation (10 000 replications) values for the $B(r,n)$ statistic.) at the $0.2, 0.15, 0.1, 0.05, 0.02$ and 0.01 levels of significance, for $r = 2$ and $n = 20, 30, \dots, 200$ in table 1.

Table 1. Critical values of $B(r,n)$ for $r=2$ (Weibull and extreme values distribution).

	α					
n	0.2	0.15	0.1	0.05	0.02	0.01
20	5.456	5.603	5.795	6.113	6.608	6.910
30	5.093	5.204	5.360	5.591	5.889	6.093
40	4.948	5.042	5.171	5.379	5.643	5.824
50	4.887	4.979	5.095	5.268	5.501	5.709
60	4.868	4.956	5.063	5.244	5.473	5.596
70	4.859	4.930	5.029	5.191	5.381	5.507
80	4.859	4.930	5.029	5.191	5.381	5.507
90	4.856	4.928	5.029	5.185	5.397	5.522
100	4.867	4.938	5.035	5.189	5.388	5.540
120	4.875	4.945	5.035	5.194	5.352	5.474
140	4.893	4.960	5.045	5.179	5.343	5.459
150	4.893	4.990	5.042	5.181	5.343	5.484
160	4.925	4.993	5.074	5.200	5.350	5.455
170	4.922	4.990	5.075	5.214	5.388	5.487
180	4.932	4.999	5.091	5.222	5.372	5.464
190	4.941	5.007	5.089	5.214	5.372	5.478
200	4.953	5.019	5.101	5.231	5.383	5.474

§6. Power comparisons by simulation

We generated the data under the alternative $H_{2,k,1}$ for various shape-scale families, various values of k. For Weibull and loglogistic families we fixed $n=100$ and for the lognormal family $n=80$.

For any of considered tests and any fixed n and k the frequence of one of the following three possible events were computed: (1) all the r largest observations are from the alternative distribution (2) all the r largest observations are from zero distribution (3) the r largest observations are from both distributions. The probabilities of these events are defined as *non-spurious power, spurious power and swamping effect*, respectively. The most important is to have good non-spurious power, i.e. identify all the contaminants. The power of the test is the sum of the three above defined. As a rule the part of the non-spurious power in the sum increases when the alternative goes away from the null hypothesis. In the case of

Table 2. Critical values of $B(r,n)$ for $r = 2$ (lognormal and normal distribution).

	α					
n	0.2	0.15	0.1	0.05	0.02	0.01
20	5.462	5.649	5.907	6.363	6.866	7.326
40	5.298	5.441	5.651	5.977	6.380	6.676
60	5.389	5.536	5.726	6..041	6.408	6.690
80	5.464	5.594	5.765	6.057	6.417	6.636
100	5.544	5.684	5.855	6.124	6.492	6.731
120	5.617	5.751	5.936	6.220	6.558	6.807
140	5.688	5.828	6.012	6.277	6.624	6.841
160	5.750	5.884	6.049	6.323	6.672	6.900
180	5.802	5.938	6.106	6.386	6.696	6.899
200	5.853	5.983	6.162	6.457	6.764	6.991

Table 3. Critical values of $B(r,n)$ for $r = 2$ (loglogistic and logistic distribution).

	α					
n	0.2	0.15	0.1	0.05	0.02	0.01
20	9.950	10.382	10.963	12.043	13.446	14.545
40	10.164	10.570	11.079	11.986	13.284	14.261
60	10.630	10.996	11.537	12.394	13.582	14.314
80	11.005	11.394	11.899	12.810	14.003	14.779
100	11.354	11.735	12.293	13.113	14.398	15.207
120	11.616	12.005	12.585	13.598	14.789	15.616
140	11.863	12.264	12.826	13.781	14.944	15.801
160	12.114	12.536	13.097	14.008	15.154	16.096
180	12.323	12.770	13.345	14.327	15.347	16.126
200	12.525	12.943	13.530	14.455	15.687	16.524

shape-scale families the alternative goes away from the null hypothesis if the parameter k increases.

We computed all types of power by simulation for various shape-scale families. All three Dixon's tests, two Grubbs tests , Tietjen–Moore (TM),

Table 4. Power of tests against $H_{r,k,1}$ (Weibull distribution) $n = 100,\ \alpha = 0.05,\ r = 2$.

	Tests								
k	Dixon-1	Dixon-2	Dixon-3	Grubbs	Grubbs-2	TM	Mann	Kumar	BZN
5	0.291	0.301	0.299	0.292	0.282	0.292	0.147	0.151	0.305
10	0.601	0.609	0.610	0.601	0.588	0.601	0.379	0.379	0.628
15	0.761	0.771	0.771	0.771	0.747	0.761	0.569	0.550	0.792
20	0.831	0.838	0.838	0.841	0.832	0.842	0.673	0.645	0.857
25	0.880	0.884	0.884	0.887	0.881	0.888	0.748	0.721	0.899
30	0.910	0.912	0.912	0.916	0.912	0.916	0.804	0.775	0.926
40	0.943	0.944	0.946	0.946	0.944	0.948	0.868	0.840	0.953
50	0.961	0.964	0.964	0.965	0.962	0.965	0.906	0.879	0.969
60	0.970	0.972	0.972	0.974	0.974	0.976	0.931	0.913	0.978

Table 5. Non-spurious power of tests against $H_{r,k,1}$ (Weibull distribution) $n = 100,\ \alpha = 0.05,\ r = 2$.

	Tests								
k	Dixon-1	Dixon-2	Dixon-3	Grubbs	Grubbs-2	TM	Mann	Kumar	BZN
5	0.075	0.077	0.077	0.090	0.089	0.090	0.039	0.037	0.096
10	0.284	0.288	0.287	0.308	0.307	0.308	0.190	0.175	0.318
15	0.448	0.452	0.452	0.471	0.467	0.470	0.345	0.317	0.479
20	0.554	0.558	0.558	0.573	0.572	0.574	0.461	0.429	0.582
25	0.630	0.631	0.631	0.644	0.643	0.644	0.550	0.515	0.651
30	0.685	0.686	0.686	0.698	0.697	0.698	0.628	0.586	0.702
40	0.750	0.751	0.751	0.756	0.766	0.767	0.701	0.674	0.760
50	0.796	0.797	0.798	0.801	0.807	0.808	0.760	0.734	0.803
60	0.828	0.829	0.829	0.833	0.838	0.838	0.804	0.784	0.835

Table 6. Spurious power of tests against $H_{r,k,1}$ (Weibull distribution) $n = 100,\ \alpha = 0.05,\ r = 2$.

	Tests								
k	Dixon-1	Dixon-2	Dixon-3	Grubbs	Grubbs-2	TM	Mann	Kumar	BZN
5	0.014	0.014	0.013	0.018	0.018	0.018	0.016	0.017	0.015
10	0.005	0.006	0.005	0.006	0.006	0.006	0.007	0.006	0.006
20	0.001	0.001	0.001	0.002	0.002	0.002	0.002	0.002	0.002
25	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
50	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Mann, Kumar and Bagdonavičius–Zerbet–Nikulin (BZN) tests were considered.

We see that when k increases then the power and the non-spurious power increases, the part of spurious power decreases quickly to zero and the part of swamping effect decreases but not so quickly as in the case of the spurious power. The same pattern we obtained in the case of the lognormal and loglogistic families so we give only power and non-spurious power for these families.

Table 7. Swamping effect of tests against $H_{r,k,1}$ (Weibull distribution) $n = 100,\ \alpha = 0.05,\ r = 2$.

	Tests								
k	Dixon-1	Dixon-2	Dixon-3	Grubbs	Grubbs-2	TM	Mann	Kumar	BZN
5	0.202	0.209	0.209	0.184	0.175	0.184	0.092	0.098	0.193
10	0.312	0.315	0.318	0.287	0.275	0.287	0.183	0.198	0.305
15	0.310	0.316	0.317	0.296	0.276	0.288	0.220	0.229	0.309
20	0.276	0.279	0.279	0.266	0.259	0.266	0.210	0.214	0.273
25	0.249	0.252	0.252	0.242	0.237	0.242	0.197	0.205	0.247
30	0.224	0.226	0.226	0.217	0.214	0.218	0.183	0.189	0.224
40	0.193	0.193	0.194	0.189	0.178	0.181	0.166	0.166	0.193
50	0.165	0.166	0.165	0.163	0.155	0.157	0.145	0.145	0.166
60	0.142	0.143	0.142	0.141	0.136	0.138	0.127	0.129	0.143

Table 8. Power of tests against $H_{r,k,1}$ (lognormal distribution) $n = 80,\ \alpha = 0.05,\ r = 2$.

	Tests								
k	Dixon-1	Dixon-2	Dixon-3	Grubbs	Grubbs-2	TM	Mann	Kumar	BZN
3	0.148	0.145	0.145	0.191	0.186	0.190	0.088	0.083	0.187
3.5	0.256	0.251	0.251	0.358	0.353	0.358	0.143	0.131	0.347
4	0.406	0.403	0.403	0.582	0.577	0.580	0.236	0.219	0.571
4.5	0.038	0.037	0.037	0.059	0.787	0.787	0.019	0.018	0.059
5	0.747	0.740	0.741	0.922	0.918	0.917	0.573	0.507	0.915
6	0.943	0.941	0.939	0.997	0.997	0.996	0.884	0.809	0.997
7	0.994	0.993	0.993	1.000	1.000	1.000	0.990	0.954	1.000

Table 9. Non-spurious power of tests against $H_{r,k,1}$ (lognormal distribution) $n = 80,\ \alpha = 0.05,\ r = 2$.

	Tests								
k	Dixon-1	Dixon-2	Dixon-3	Grubbs	Grubbs-2	TM	Mann	Kumar	BZN
3	0.038	0.037	0.037	0.059	0.057	0.057	0.019	0.018	0.059
3.5	0.115	0.112	0.112	0.173	0.173	0.172	0.060	0.053	0.170
4	0.260	0.258	0.259	0.381	0.384	0.383	0.151	0.135	0.377
4.5	0.466	0.462	0.460	0.623	0.627	0.623	0.315	0.273	0.619
5	0.673	0.668	0.668	0.817	0.818	0.816	0.518	0.454	0.812
6	0.931	0.930	0.928	0.981	0.979	0.979	0.872	0.798	0.981
7	0.994	0.993	0.992	0.999	0.999	0.999	0.989	0.953	0.999

Similar results were obtained for the alternatives $H_{2,1,\Delta}$ for various values of Δ and also for $r = 3$ but we did not include them for the lack of space.

Table 10. Power of tests against $H_{r,k,1}$ (loglogistic distribution) $n = 100,\ \alpha = 0.05,\ r = 2$.

k	Tests								
	Dixon-1	Dixon-2	Dixon-3	Grubbs	Grubbs-2	TM	Mann	Kumar	BZN
20	0.048	0.047	0.046	0.060	0.063	0.061	0.052	0.054	0.062
50	0.053	0.053	0.051	0.088	0.091	0.087	0.060	0.063	0.094
100	0.066	0.065	0.063	0.129	0.136	0.120	0.076	0.072	0.145
200	0.095	0.094	0.091	0.203	0.234	0.210	0.096	0.092	0.227
300	0.119	0.121	0.118	0.280	0.307	0.281	0.112	0.103	0.309
500	0.161	0.158	0.157	0.378	0.428	0.391	0.156	0.146	0.417
800	0.212	0.208	0.208	0.500	0.538	0.501	0.203	0.185	0.540
1500	0.295	0.290	0.284	0.665	0.706	0.667	0.289	0.262	0.703
2000	0.347	0.344	0.342	0.742	0.774	0.744	0.332	0.298	0.774
2500	0.384	0.378	0.372	0.793	0.817	0.796	0.377	0.351	0.820
5000	0.493	0.487	0.485	0.899	0.915	0.900	0.501	0.457	0.915

Table 11. Non-spurious power of tests against $H_{r,k,1}$ (loglogistic distribution) $n = 100,\ \alpha = 0.05,\ r = 2$.

k	Tests								
	Dixon-1	Dixon-2	Dixon-3	Grubbs	Grubbs-2	TM	Mann	Kumar	BZN
20	0.002	0.002	0.002	0.002	0.003	0.002	0.002	0.002	0.003
50	0.009	0.009	0.009	0.016	0.016	0.015	0.009	0.009	0.016
100	0.020	0.020	0.020	0.043	0.045	0.040	0.019	0.018	0.047
200	0.047	0.047	0.045	0.101	0.122	0.108	0.040	0.038	0.110
300	0.075	0.076	0.075	0.169	0.186	0.170	0.059	0.054	0.183
500	0.115	0.114	0.114	0.255	0.297	0.273	0.096	0.087	0.275
800	0.169	0.166	0.167	0.373	0.400	0.374	0.142	0.128	0.396
1500	0.255	0.252	0.247	0.540	0.575	0.541	0.228	0.207	0.568
2000	0.310	0.308	0.307	0.627	0.652	0.628	0.276	0.250	0.647
2500	0.348	0.344	0.372	0.684	0.704	0.686	0.323	0.302	0.702
5000	0.469	0.463	0.461	0.817	0.831	0.818	0.456	0.419	0.839

§7. Conclusion

The power analysis shows that for all considered families of distributions the Dixon-1, Dixon-2, Dixon-3, Mann and Kumar tests have considerably smaller power and non-spurious power than the new (BZN), Grubbs, Grubbs-2 and Tietjen–Moore tests.

For the Weibull distribution the best three tests in terms of non-spurious power are:

For alternatives $k = 5, 10, 15, 20, 25, 30$: 1. BZN; 2–3. Grubbs and TM.
For alternatives $k = 40, 50$: 1. TM; 2. Grubbs-2; 3. BZM.
For alternatives $k = 60$: 1. TM; 2. BZM; 3. Grubbs.

So the new test is a very strong competitor to TM and Grubbs tests, especially against near alternatives. Other tests are weaker.

For the loglogistic distribution the best three tests in terms of non-spurious power are:

For alternatives $k = 20, 50$: BZN. Grubbs and Grubbs-2 tests are almost equivalent;
For alternatives $k = 100$: 1. BZM; 2. Grubbs-2; 3. Grubbs.
For alternatives $k = 200, 300, 500, 800$: 1. Grubbs-2; 2. BZM; 3. TM.
For alternatives $k = 1500, 2000, 2500$: 1. Grubbs-2; 2. BZM; 3. TM.
For alternatives $k = 5000$: 1. BZM; 2. Grubbs-2; 3. TM.

So the new test is a strong competitor to Grubbs-2 test.

For the lognormal distribution the best three tests in terms of non-spurious power are:

For alternatives $k = 3$: 1–2. BZN and Grubbs; 2. Grubbs-2; 3. BZM.
For alternatives $k = 3.5, 4, 4.5, 5$: 1–2. Grubbs and Grubbs-2; 3. TM;
For alternatives $k = 6$: 1–2. Grubbs and BZN; 3–4. Grubbs-2 and TM.
For alternatives $k = 7$: Grubbs, Grubbs-2, BZN and TM tests are equivalent.

So the new test is not a bad competitor to Grubbs tests. Other tets are weaker.

Concluding, we recommend as the most powerfull the BZN, TM and Grubbs tests for the Weibull distribution, Grubbs-2 and BZN tests for the loglogistic distribution, Grubbs, Grubbs-2 and BZN tests for lognormal distribution

References

1. N. Balakrishnan, *Permanents, order statistics, outliers, and robustness.* — Revista Matemática Complutense **20** (2007), 7–107.
2. V. Barnett, T. Lewis, *Outliers in Statistical Data.* Wiley, New York, 1994.
3. Л. Н. Большев, М. К. Убайдулаева, *Критерий Шовенэ в классической теории ошибок.* — Теория вероятн. и ее примен. **19** (1974), 683–692.
4. Л. Н. Большев, Н. В. Смирнов, *Таблицы математической статистики.* ВЦ АН СССР, М., 1968.
5. M. S. Chikkagoudar, S. H. Kunchur, *Distribution of test statistics for multiple outliers in exponential samples.* — Commun. Statist. Theory Methods **12** (1983), 2127–2142.
6. W. J. Dixon, *Analysis of extreme values.* — Ann. Math. Statist. **21** (1950), 488–506.

7. K. Y. Fung, S. R. Paul, *Comparisons of outlier detection procedures in Weibull or extreme-value distribution.* — Commun. Statist. - Simulation Computation **14** (1985), 895–917.
8. F. E. Grubbs, *Sample criteria for testing outlying observations.* — Ann. Math. Statist. **21** (1950), 27–58.
9. И. А. Ибрагимов, Н. М. Халфина, *Некоторые асимптотические результаты, связанные с критерием Шовенэ.* — Теория вероятн. и ее примен. **23**, 3 (1978), 615–619.
10. D. G. Kabe, *Testing outliers from an exponential population.* — Metrika **15** (1970), 15–18.
11. N. Kumar, *Test for multiple upper outliers in an exponential sample irrespective of origin.* — Statistics **47** (2013), 184–190.
12. S. Lalitha, N. Kumar, *Testing for upper outliers in gamma sample.* — Commun. Statist. - Theory Methods **41** (2012), 820–828.
13. T. Lewis, N. R. Fiellerm, *A recursive algorithm for null distribution for outliers: I. gamma samples.* — Technometrics **21** (1979), 371–376.
14. J. Likes, *Distribution of Dixon's statistics in the case of an exponential population.* — Metrika **11** (1966) 46–54.
15. Ю. В. Линник, *Метод наименьших квадратов и основы теории обработки наблюдений.* — Физматлит, М., 1962.
16. N. R. Mann, E. M. Scheuer, K. W. Fertig, *A new goodness-of-fit test for the two parameter Weibull or extreme-value distribution with unknown parameters.* — Commun. Statist. - Theory Methods **2** (1973), 383–400.
17. N. R. Mann, *Optimal outlier tests for a Weibull model to identify process changes or predict failure times.* — Studies in the Management Sciences **19** (1981), 261–270.
18. В. И. Пагурова, *О критерии Шовенэ для обнаружения нескольких выбросов.* — Теория вероятн. и ее примен. **30**, 3 (1985), 558–561.
19. R. Pyke, *Spacings.* — J. Roy. Statist. Soc. **3** (1965), 395–449.
20. B. Rosner, *On the detection of many outliers.* — Technometrics **17** (1975), 221–227.
21. G. Tietjen, R. Moore, *Some Grubbs-type statistics for the detection of several outliers.* — Technometrics **14** (1972), 583–597.
22. A. Zerbet, M. Nikulin, *A new statistic for detecting outliers in exponential case.* — Commun. Statist. – Theory Methods **32** (2003), 573–584.

Department of Mathematical Statistics,
University of Vilnius, Lithuania

Поступило 2 ноября 2015 г.

UFR Siences and Modelization,
University of Bordeaux,
Bordeaux, France
E-mail: mikhail.nikouline@free.fr

FSJES BP 8658
Poste Dakhla Agadir,
Maroc