ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 152, 2000

*Jerzy Korzeniewski**

SAMPLE BREAKDOWN POINTS OF THE WILCOXON AND SIGN TESTS FOR LOCATION

Abstract. In 1996 Zhang introduced sample replacement points for the level and power of tests and their simplified versions. This paper presents numerical values of the breakdown points of the Wilcoxon and sign tests for location for the normal distribution. The results confirm the conclusions of He *et al.* (1990) as well as the asymptotic dominance of the power breakdown points of the sign tests over the Wilcoxon test. The breakdown points of the acceptance decision show a bit different behaviour.

1. INTRODUCTION

Let us supose that a certain random variable takes values in space (**X**, **A**). Let **M** be the set of all probability measures on this space. The random sample $x_1, ..., x_n$ will be denoted by X or X_n; $\varphi(X)$ – the value of test decision function; $X_k = (x_1, ..., x_k)$ – first k components of sample X; $Y_k \cup Z_l$ denotes sample $y_1, ..., y_k, z_1, ..., z_l$ and Y_k denotes sample $y_1, ..., y_k$, Z_l – is sample $z_1, ..., z_l$; $N = \{1, 2, ..., n\}$ and for $I_k = \{i_1, ..., i_k\} \subset \subset N$, $1 \leqslant k \leqslant n$ let $\mathbf{X}(I_k) = \{Y = (y_1, ..., y_n)\colon\ y_i = x_i$ for $i \in I_k$, $y_j \in R$ for $j \notin I_k\}$.

Let us recall the following definitions introduced by Zhang (1996).

Definition 1. The replacement sample breakdown point of the acceptance decision of φ at X is given by

$$\varepsilon_{RA} = \frac{1}{n} \min\{m : n \geqslant m \geqslant 0, \min_{I_{n-m} \in N} \sup_{Y \in \mathbf{X}(I_{n-m})} \varphi(Y) = 1\}.$$

* University of Łódź, Chair of Statistical Methods.

The replacement sample breakdown point of the rejection decision of φ at X defined by

$$\varepsilon_{RR} = \frac{1}{n} \min\{m : n \geqslant m \geqslant 0, \min_{I_{n-m} \in N} \sup_{Y \in X(I_{n-m})} \varphi(Y) = 0\}.$$

These breakdown points represent the smallest percent of the worst possible contamination which causes (independent of other observations) a given decision. These breakdown points are calculated for a given sample X. The calculations sometimes are difficult for large sample sizes. To measure the behaviour of the test function φ we may use the simplified replacement sample breakdown points which also depend on a sample but as we will see can be used with respect to the set function.

Definition 2. The simplified replacement sample breakdown points of the acceptance decision and rejection decision are given by

$$\varepsilon_{RA}(X) = \frac{1}{n} \min\{m : n \geqslant m \geqslant 0, \sup_{Y \in R^m} \varphi(X_{n-m} \cup Y) = 1\},$$

$$\varepsilon_{SR}(X) = \frac{1}{n} \min\{m : n \geqslant m \geqslant 0, \inf_{Y \in R^m} \varphi(X_{n-m} \cup Y) = 0\}$$

respectively.

One of the ways to apply the idea of breakdown points is to assess sample reliability. For example (compare Zhang, 1993) let us consider the following sample X: 1.2, 2.4, 1.3, 1.3, 0.0, 1.8, 0.8, 4.6, 1.4 from random variable with distribution $N(\mu, \sigma)$. We want to test the null hypothesis $H_0 : \mu = 1.4$ against the alternative $H_1 : \mu \neq 1.4$. If we use the two-sided t-test with critical function of the form

$$\varphi(X) = \begin{cases} 1, & \text{if } T(X)^2 > c_n, \\ 0, & \text{otherwise,} \end{cases}$$

where $T(X) = \overline{X}/S(X)$, $S(X) = \sqrt{\sum(x_i - \overline{X})^2/(n-1)}$, we find that $T(X)^2 < 0.51166$ and we accept the null hypotheses at the 0.05 level of significance. There is one outlying value in sample X namely 4.6. If we are worried about the validity of the decision we may calculate the value of the sample acceptance decision breakdown point. We get $\varepsilon_{SA}(X) = 2$ which suggests that at least two observation errors are needed to change the decision we made. Since 4.6 is the only possible error therefore for the sample X we do not need any more robust test. This application of the concept of breakdown points is connected with data analysis.

Another possible way of application of this idea is to assume that the test function φ is permutation invariant with respect to sample observations and to use simplified breakdown points to measure the breakdown robustness of tests. In this sense we can use the probabilities

$$P(\varepsilon_{SR}(X) \geqslant w \mid \varphi(X) = 1), \quad P(\varepsilon_{SA}(X) \geqslant w \mid \varphi(X) = 0),$$

because for any $w \geqslant 0$

$$P(\varepsilon_{SR}(X) \geqslant w \mid \varphi(X) = 1) = P(\varepsilon_{SR}(X^{*}) \geqslant w \mid \varphi(X^{*}) = 1),$$

$$P(\varepsilon_{SA}(X) \geqslant w \mid \varphi(X) = 1) = P(\varepsilon_{SA}(X^{*}) \geqslant w \mid \varphi(X^{*}) = 0),$$

where X^{*} is a permutation of X.

2. EXPERIMENT DESCRIPTION

For the normal distribution with unit variance and appropriate location θ we generate 30 000 samples. For every sample we calculate how big is the smallest number of observations (counting from the sample beginning) which have to be changed (in the least favourable way) to reverse the rejection decision (for the rejection decision breakdown) or the acceptance decision (for the acceptance decision breakdown). For the acceptance decision breakdown points we may use any other distribution (e.g. uniform) because both tests are nonparametric and are independent of distribution under the null hypotheses.

The hypotheses tested are the following

$$H_0 : \theta = 0, \quad H_1 : \theta > 0.$$

We test these hypotheses with the help of the sign and the Wilcoxon test. Both tests are randomized so that the level of significance is equal 0.05 (for tests description see Domański, 1990).

3. CONCLUSIONS

Looking at the results we can compare the rejection decision robustness and the acceptance decision robustness of both tests. The sign test is more robust than the Wilcoxon test as far as the rejection robustness (tab. 2) is concerned at every θ considered i.e. 0.1, 0.2, 0.4, 0.6, 0.8, 1. This conclusion is in accordance with the comparison of the power breakdown functions given by He *et al.* (1990). He defines the power breakdown function of statistic T at distribution F_θ, $\theta \in H_1$ as

$$\varepsilon^*(F_\theta) = \inf\{\varepsilon > 0 : T_{H_0} \cap T((1-\varepsilon)F_\theta + \varepsilon G) \neq 0 \quad \text{for some } G\}.$$

The rejection decision breakdown point seems to be the finite sample version of the power breakdown function, therefore the conclusions are not surprising. The behaviour of the acceptance decision breakdown points (tab. 1, $\theta = 0$), however, is a bit different – the Wilcoxon test looks a little better. If there is no probability at a given value of breakdown points (and at higher values) it means that theoretically, there is no possibility of appearing of this value. The asymptotic behaviour of breakdown points for certain classes of tests is given by Zhang (1997) therefore analysing larger sample sizes is not necessary.

Table 1

Probabilities of the simplified acceptance decision breakdown points

ε_{RA}	Sign test			Wilcoxon test		
	$n = 10$	$n = 20$	$n = 30$	$n = 10$	$n = 20$	$n = 30$
1	.0354	.0251	.0190	.0313	.0213	.0168
2	.0553	.0329	.0255	.0458	.0294	.0198
3	.0790	.0431	.0336	.0673	.0374	.0271
4	.1160	.0582	.0439	.0915	.0479	.0328
5	.1526	.0727	.0506	.1266	.0600	.0392
6	.1903	.0918	.0626	.1562	.0699	.0493
7	.2014	.1062	.0718	.1808	.0847	.0543
8	.1437	.1148	.0802	.1787	.0943	.0628
9	.0262	.1178	.0912	.1218	.1031	.0734
10		.1126	.0938		.1048	.0776
11		.0948	.0921		.0985	0.799
12		.0708	.0862		.0875	.0776
13		.0398	.0776		.0682	.0792
14		.0160	.0639		0.502	.0733
15		.0033	.0460		.0290	.0644
16			.0321		.0107	.0556
17			.0182		.0025	.0436
18			.0085		.0002	.0321
19			.0025			.0122
20			.0006			.0193
21						.0059
22						.0028
23						.0008
24						.0001

Table 2

Probabilities of the simplified rejection decision breakdown points

ε_{RR}	Sign test						Wilcoxon test					
$\theta =$	0.1	0.2	0.4	0.6	0.8	1	0.1	0.2	0.4	0.6	0.8	1
$n = 10$												
1	.6260	.5995	.5454	.4705	.3958	.3077	.946	.941	.922	.891	.833	.758
2	.2964	.3131	.3354	.3681	.3891	.3980	.054	.059	.078	.109	.167	.252
3	.0777	.0875	.1192	.1614	.2151	.2943						
$n = 20$												
1	.4202	.3849	.2974	.2052	.1128	.0489	.7150	.6609	.5171	.3310	.1488	.0385
2	.2831	.2812	.2625	.2216	.1603	.0943	.2434	.2800	.3570	.4166	.3685	.2257
3	.1661	.1774	.2025	.2137	.1988	.1527	.0402	.0567	.1189	.2331	.4230	.5803
4	.0850	.0982	.1351	.1762	.2140	.2151	.0001	.0002	.0007	.0194	.0597	.1554
5	.0321	.0419	.0711	.1182	.1784	.2355						
6	.0098	.0136	.0258	.0523	.1022	.1819						
7	.0017	.0029	.0056	.0129	.0334	.0717						
$n = 30$												
1	.3323	.2863	.1914	.0977	.0341	.0070	.5864	.5095	.3153	.1197	.0218	.0012
2	.2475	.2371	.1883	.1145	.0530	.0140	.2862	.3091	.3168	.2247	.0787	.0115
3	.1767	.1828	.1759	.1416	.0790	.0278	.0999	.1361	.2340	.2963	.2080	.0646
4	.1147	.1285	.1509	.1532	.1081	.0541	.0249	.0406	.1107	.2595	.3810	.2841
5	.0676	.0797	.1175	.1530	.1447	.0920	.0025	.0047	.0225	.0952	.2869	.5465
6	.0360	.0485	.0851	.1296	.1617	.1305	.0001	.0001	.0007	.0046	.0237	.0921
7	.0162	.0231	.0505	.1014	.1569	.1746						
8	.0064	.0102	.0261	.0632	.1292	.1949						
9	.0021	.0030	.0104	.0313	.0833	.1690						
10	.0006	.0008	.0032	.0117	.0393	.1008						
11		.0001	.0006	.0027	.0102	.0339						
12				.0001	.0005	.0015						

REFERENCES

Domański Cz. (1990), *Testy statystyczne*, PWE, Warszawa.

He X., Simpson D. G., Portnoy S. L. (1990), *Breakdown Robustness of Tests*, J. Am. Statist. Assoc., **85**, 446–452.

Zhang I. (1996), *The Sample Breakdown Points of Tests*, J. Stat. Planning and Inference, **52**, 161–181.

Zhang I. (1997), *The Optimal Breakdown M-test and Score Test*, "Statistics", **30**, 47–68.