On Power of Randomness and of Absence of Trend Tests in Dispersion Characteristics

Boris Yu. Lemeshko, Irina V. Veretel`nikova
Novosibirsk State Technical University, Novosibirsk, Russia

Abstract — **Distributions of various parametric and nonparametric tests are studied using methods of statistical modeling. Such tests are designed to test hypotheses for randomness or absence of a trend in dispersion characteristics. Procedure of interactive simulation of distributions of the test statistics is proposed and implemented. Such procedure allows valid usage of corresponding tests in case of violation of standard assumptions. The results of comparative analysis of the power of the tests against competing hypotheses with different models of linear, periodic, and mixed trend in the distribution characteristics are shown. Conclusions on preferences for using different criteria are made.**

Kay words – **Power of a test, trend, hypothesis of randomness, statistical modeling**

I. INTRODUCTION

A VARIETY of parametric and nonparametric tests has been proposed at different times to test the hypothesis for randomness or absence of a trend in the mathematical expectation and in the dispersion characteristics. However, available sources do not allow us to judge the benefits of a particular test and do not contain any distinct recommendations on the area of application and prerequisites providing correctness of statistical conclusions when using the tests under consideration. As a rule, assumption of normal distribution law of noise is the main prerequisite for ensuring the correct application of parametric tests, but it is not always realized in practice. The usage of nonparametric tests is based on asymptotic distribution of statistics of such tests. For limited sample number distribution of the statistics of parametric and non-parametric tests may differ significantly from the corresponding limit distributions of statistics used for testing the hypothesis. In case of nonparametric tests the problem gets even worse because of apparent discreteness of the statistic. In such situations, the usage of the limiting (asymptotic) distribution of the statistics instead of the true distribution of such statistics may lead to wrong conclusion.

In this paper methods of statistical modeling are used to investigate the distributions and powers of statistical tests designed to check the hypothesis against the absence of a trend in dispersion characteristics (in variance) of observed random number sequence (measured values).

II. STATEMENT OF THE TASK

When checking the absence of a trend in the dispersion characteristics the task is formulated as follows. It is assumed that time series $x_1, x_2, ..., x_n$ of mutually independent random variables is observed. The hypothesis $H_0 : \sigma_i = \sigma,\ i = 1, 2, \dots, n$,, is tested that all sample values belong to the same population with mean σ, against a competitive hypothesis about the presence of a trend $H_l : |\sigma_{i+1} - \sigma_i| > 0,\ i = 1, 2, \dots, n-1$.

When checking the absence of a variance shift (in dispersion characteristics) it is assumed that the observed sequence of measured values $x_1, x_2, ..., x_n$ has one mean μ. The hypothesis H_0: $\sigma_1^2 = ... = \sigma_n^2 = \sigma_0^2$ (σ_0^2 being unknown) is tested against competitive hypothesis

$$H_l:\ \sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2 = \sigma_0^2;$$
$$\sigma_{k+1}^2 = \sigma_{k+2}^2 = ... = \sigma_n^2 = \sigma_0^2 + \delta;\quad (\delta > 0),$$

for variance value changes in some unknown point, i.e. k is unknown ($1 \le k \le n-1$).

III. STATISTICAL TESTS

A. Cox-Stuart test

Cox-Stuart test against the hypothesis of absence of a trend in variance (in dispersion characteristics) is designed as follows.

Initial sample $x_1, ..., x_n$ is divided into $[n/k]$ subsamples with k number of elements $x_1, ..., x_k$; $x_{k+1}, ..., x_{2k}$; $x_{2k+1}, ..., x_{3k}$; ...; $x_{n-k+1}, ..., x_n$ (if n is not divided by k, then the required number of measurements in the center is dropped out). For every i th subsample the range w_i is found ($1 \le i \le r$, $r = [n/k]$). Next, the resulting sequence of ranges w_i is tested against the trend in the mean values using the test with statistics

$$S_1^* = \frac{S_1 - E[S_1]}{\sqrt{D[S_1]}}, \tag{1}$$

where

$$S_1 = \sum_{i=1}^{[n/2]} (n - 2i + 1) h_{i,n-i+1},$$

$$E[S_1] = \frac{n^2}{8},\ D[S_1] = \frac{n(n^2-1)}{24},$$

where $h_{i,j} = 1$, if $x_i > x_j$, and $h_{i,j} = 0$, if $x_i \le x_j$ ($i < j$). If the hypothesis for the absence of a trend is true, distribution (1) can be approximately described by standard normal law.

It is recommended to choose the value of k in [5] according to the following correlations:

$$n \geq 90 \rightarrow k = 5\,;\; 64 \leq n < 90 \rightarrow k = 4\,;$$

$$48 \leq n < 64 \rightarrow k = 3\,;\; n < 48 \rightarrow k = 2\,.$$

The discreteness of the S_1^* statistics distribution upon detection of a trend in the variance is significantly higher than the discreteness of the Cox–Stuart statistics distribution for trend in mean (see Fig. 1). This is natural because the analyzed range sample contains only $[n/k]$ number of elements.

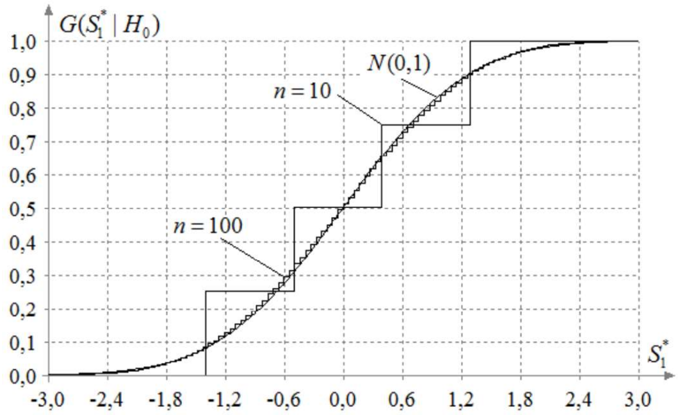

Fig. 1. Convergence to the standard normal law of the statistic distribution (1) of Cox-Stuart test

When using the Cox–Stuart test for detection of a trend in the dispersion, the difference of statistics discrete distribution from the standard normal law can almost be neglected only for $n > 170$ [2].

B. Foster-Stuart test

This nonparametric test can be used to test hypotheses of absence of a trend in the mean values or in the variances (dispersion characteristics) depending on the used statistics type. Test against trend in distribution characteristics is given by [1]:

$$S = \sum_{i=2}^{n} S_i\,, \tag{2}$$

where $S_i = u_i + l_i$;

$u_i = 1$, if $x_i > x_{i-1},\ x_{i-2},\ ...,\ x_1$, otherwise $u_i = 0$;

$l_i = 1$, if $x_i < x_{i-1},\ x_{i-2},\ ...,\ x_1$, otherwise $l_i = 0$.

It is clear that $0 \leq S \leq n-1$.

In the absence of a trend the normalized statistics

$$\tilde{t} = \frac{S-\mu}{\hat{\sigma}_S}\,, \tag{3}$$

where

$$\mu = 2\sum_{i=2}^{n}\frac{1}{i},\ \hat{\sigma}_S = \sqrt{\mu - 4\sum_{i=2}^{n}\frac{1}{i^2}} \approx \sqrt{2\ln n - 3.4253}$$

are approximately described by Student's distribution with $\nu = n$ degrees of freedom. The hypothesis of absence of a trend is rejected at large modulus values of statistics (3).

Actually, the area of discrete values is the range of definition of $\tilde{t}$ statistics. The analysis of statistics distributions shows that even with relatively large sample sizes (around $n = 100,\ 200$) the discrete distributions of test statistics are significantly different from the Student distribution with n degrees of freedom [3]. Statistics distribution functions $\tilde{t}$ are shown Fig. 2.

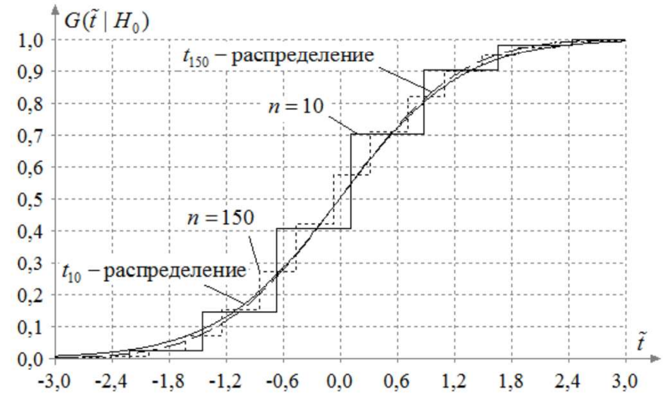

Fig. 2. Statistic distribution functions (3) of Foster-Stewart test in dependence of sample sizes.

This implies that usage of reached significance point (*p-value*) instead of real (discrete) statistic distributions of Student asymptotic t-distributions could lead to sensible errors.

C. Hsu test against variance shift and shift point detecting

Under this test the rejection of the hypothesis of randomness (for absence of a trend) can show the discovery of a *variance shift*. Hsu test statistics are given by [5]

$$H = \frac{\sum_{i=1}^{n}(i-1)(x_i - m_x)^2}{(n-1)\sum_{i=1}^{n}(x_i - m_x)^2},\ 0 \leq H \leq 1\,, \tag{4}$$

where m_x is median of variation series. Under the assumption that the mathematical expectation of a sequence of random variables has the same value, the hypothesis of a constant variances is tested. As a competitive hypothesis, the change in the dispersion of observed values at some (unknown) time (starting from some element of the sample) can be considered. The test is two-sided: the tested hypothesis of absence of a variance shift is rejected for small and large values of the statistics (4).

Usually the test is used in a normalized form

$$H^* = \frac{H - 1/2}{\sqrt{D[H]}}, \tag{5}$$

where $D[H] = \dfrac{n+1}{6(n-1)(n+2)}$.

Under the validity of the hypothesis of the absence of variance changes in the asymptotic, such statistics (5) obeys the standard normal law.

The modeling results [3] show that for $n > 30$ statistics distribution agrees well with the standard normal law.

Hsu test is a parametric test. Likewise any other parametric test dealing with variance hypotheses, distribution of its statistic significantly depends on the law to which the analyzed random values belong (noise distribution law). Statistics obtained using distribution simulation (5) when random values belong to summary generic law (6) for different values of form parameters are shown in Fig. 3.

$$f(x) = \frac{\theta_2}{2\theta_1\Gamma(1/\theta_2)}\exp\left(-\left(\frac{|x-\theta_0|}{\theta_1}\right)^{\theta_2}\right) \tag{6}$$

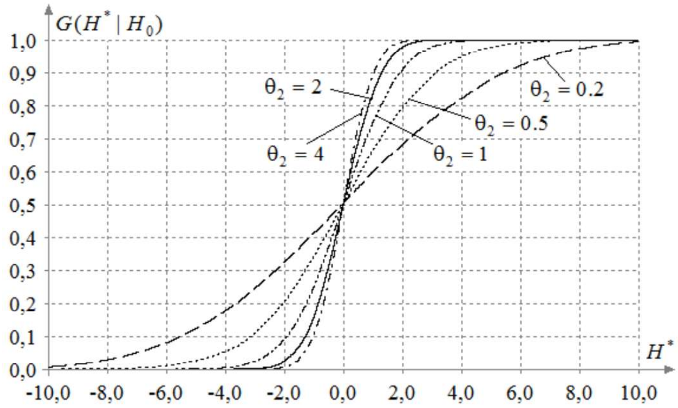

Fig. 3. Statistic distributions (5) when random values belong to family laws (6) for different values of form for *n=100*

Statistics distribution (5) strongly depends on the law of distribution to which random variables belong. The greatest deviation from the standard normal law is observed in the case when random variables belong to the laws with heavy tails. Asymmetry of the law significantly affects the statistics distribution.

A test allowing to determine the change point of the variance (in the case when observations belong to the normal law) is proposed in [5]. Statistics of this test are presented as follows. Let for $k = 1, 2, ..., n-1$

$$w_k = \sum_{i=1}^{k} (x_i - m_x)^2 \,,\; W_k = \frac{w_n - w_k}{w_k} \frac{k}{n-k},$$

where k corresponds to the required variance change point. If x_i belongs to normal law, then values of W_k, $k = 1, 2, ..., n-1$, belong to corresponding $F_{n-k,k}(W)$ Fisher distributions with $n-k$ and k degrees of freedom.

Next, based on the corresponding distribution functions, we find $\gamma_k = F_{n-k,k}(W_k)$, where γ_k must obey to uniform law under the absence of *variance shift*.

G-test statistics are given by

$$G = \frac{1}{n-1} \sum_{k=1}^{n-1} \gamma_k \,,\; 0 \le G \le 1 \,. \tag{7}$$

The hypothesis about absence of variance changes is rejected with significance level α, if $G < G_{\alpha/2}$ or $G > G_{1-\alpha/2}$. In this case value k corresponding to the maximum value $|\gamma_k - 1/2|$, evaluates the desired change point of the variance value in observed series. For $x_1 = m_x$ value $w_1 = 0$, thus $W_1 = \infty$ and $\gamma_1 = 1$.

The type of limit distribution of the statistics (7) is not given in the original material, only percentage points are given.

Based on the results of the statistical modeling we have shown that a good model of the limit distribution of the statistics (7) is a beta distribution of the 1st kind with the density of

$$f(x) = \frac{1}{\theta_2 \mathrm{B}(\theta_0, \theta_1)} \left(\frac{x-\theta_3}{\theta_2} \right)^{\theta_0 - 1} \left(1 - \frac{x-\theta_3}{\theta_2} \right)^{\theta_1 - 1}$$

and parameter values $\theta_0 = 2.7663$, $\theta_1 = 2.7663$, $\theta_2 = 1$, $\theta_3 = 0$ (see Fig. 4). Based on this law we can find percentage points $G_{\alpha/2}$ and $G_{1-\alpha/2}$ or *p*-values.

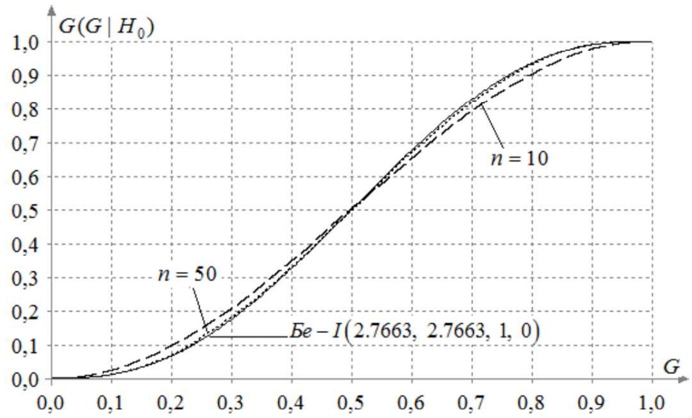

Fig. 4. Convergence to type one beta-distribution of the statistic distributions (10)

G-test is a parametric test as well, so distribution of its statistic significantly depends on the type of the observed law.

In case of violation of standard assumption of x_i normality by the value laws W_k, $k = 1, 2, ..., n-1$, would not be the Fisher distributions. For the correct application of test in non-standard conditions calculating the actual distribution of W_k, $k = 1, 2, ..., n-1$ values must be found and used for calculations of γ_k (which is preferable because the distribution of the G-statistic will be the same), or one can use the Fisher distribution. But in these conditions there must be an unknown distribution of G-statistic (which is less preferable but easier to implement, because only one distribution is to be found.)

D. Klotz and Savage rank tests against variance shifts

Rank tests for detecting the change of the scale parameter (dispersion characteristic) in the unknown point are based on the usage of a family of rank statistics in form [6]

$$S_R = \sum_{i=1}^{n} i a_n(R_i), \tag{8}$$

where R_i are ranks of sampled values in an ordered series of measurements.

Tests differ by the used scores a_n. Their type determines the name of the test. The following scores are commonly used:

- *Klotz scores* $a_{1n}(i) = U^2_{i/(n+1)}$, where U_γ – is a γ-quantile of standard normal law;
- *Savage scores* $a_{2n}(i) = \sum_{j=1}^{i} \frac{1}{n-j+1}$.

If the tested hypothesis H_0 is true, then tests with statistics $S_{R,j} = \sum_{i=1}^{n} i a_{jn}(R_i)$, $j = 1, 2$ are free from the distribution and are symmetric with respect to $E[S_{R,j}] = \frac{n+1}{2} \sum_{i=1}^{n} a_{jn}(i)$.

Usually normalized tests with the following statistics are used

$$S^*_{R,j} = \frac{S_{R,j} - E[S_{R,j}]}{\sqrt{D[S_{R,j}]}}, \tag{9}$$

where

$$E[S_{R,1}] = \frac{n+1}{2} \sum_{i=1}^{n} U^2_{i/(n+1)} \,,\; E[S_{R,2}] = \frac{n(n+1)}{2};$$

$$D[S_{R,1}] = \frac{n(n+1)}{12}\sum_{i=1}^{n} U^4_{i/(n+1)} - \frac{1}{3n+3}\left[E[S_{R,1}]\right]^2$$

$$D[S_{R,2}] = \frac{n(n+1)}{12}(n - \sum_{j=1}^{n}\frac{1}{j}).$$

Statistics (8) are approximately obeying the standard law. The convergence of the statistics distributions to the standard law was studied in [**Ошибка! Источник ссылки не найден.Ошибка! Источник ссылки не найден.**].

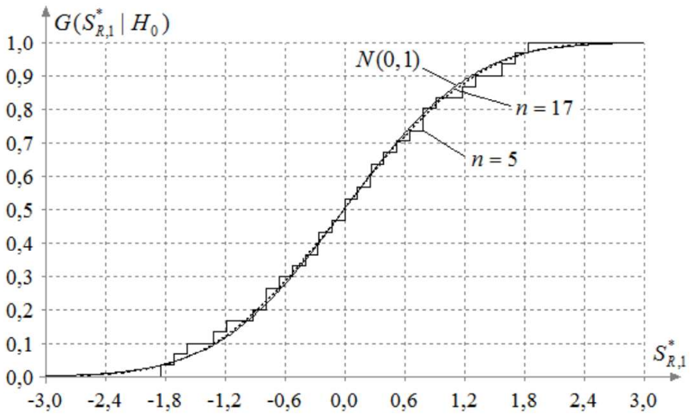

Fig. 5. Convergence to standard normal law of $S^*_{R,1}$ statistic distribution of the rank test with Klotz scores

Statistical modeling research of the distribution of statistics with Klotz scores had shown that for $n > 20$ distribution is well-approximated by standard normal law (see Fig. 5). Distribution of the test statistics with Savage scores is also matching well with standard normal law, but only for $n > 30$ (see Fig. 6).

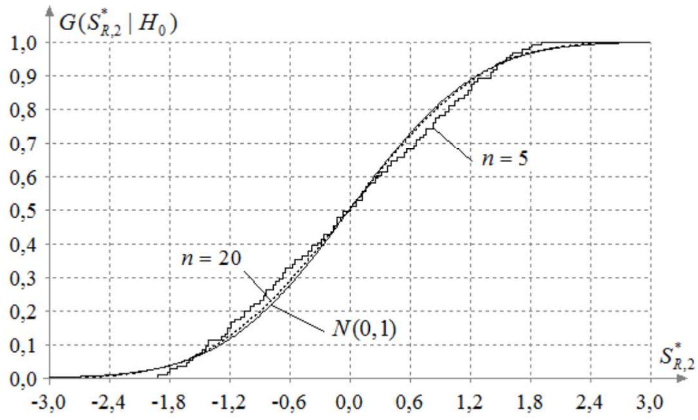

Fig. 6. Convergence to standard normal law of $S^*_{R,2}$ statistic distribution of the rank test with Savage scores

IV. DISCUSSION OF RESULTS

Analysis of the power of the tests was carried out for the situation when observed random variables belong to the normal law. The tested hypothesis H_0 corresponds to true assumption on independency of observed random values (in other words, the absence of trend). Various situations corresponding to the presence of a trend in the variance were used as competing hypotheses.

During analysis of test powers for the tests against variance change in an unknown point hypotheses close to the H_0 (in case of normal distribution of random variables) were treated as competitive, when at some point the standard deviation was increased by 5, 10, 15%:

$$H_1:\ \sigma_1^2 = ... = \sigma_k^2 = 1;\ \ \sigma_{k+1}^2 = ... = \sigma_n^2 = 1.1025,$$

$$H_2:\ \sigma_1^2 = ... = \sigma_k^2 = 1;\ \ \sigma_{k+1}^2 = ... = \sigma_n^2 = 1.21,$$

$$H_3:\ \sigma_1^2 = ... = \sigma_k^2 = 1;\ \ \sigma_{k+1}^2 = ... = \sigma_n^2 = 1.3225,$$

where $k = n/2$.

One competitive hypothesis was considered as more distant

$$H_4:\ \sigma_1^2 = ... = \sigma_k^2 = 1;\ \ \sigma_{k+1}^2 = ... = \sigma_n^2 = 4.$$

As the time series that correspond to hypothesis H_1 are visually similar to the time series corresponding to hypotheses H_2, H_3, Fig. 7 shows only time series for H_1.

The presence of a linear trend in the dispersion characteristics of the observed series of random variables (change in scale parameter) in the interval $t \in [0,1]$ can be simulated according to

$$x_i = \xi_i(1 + ct_i), \tag{10}$$

where $c \in (-1, \infty)$, $t_i = (i-1)\Delta t$, $\Delta t = 1/n$. True tested hypothesis H_0 corresponds to parameter value $c = 0$.

In case of a periodic trend in the characteristics of dispersion, random values can be simulated, for example, in accordance with the following formula

$$x_i = \xi_i(1 + d \cdot \sin(2k\pi t_i))$$

for $|d| < 1$. In case of a combined trend it can be simulated according to

$$x_i = \xi_i(1 + ct_i + d\sin(2k\pi t_i)),$$

for $|d| < 1$, if $c \geq 0$, and for $|d| < 1 + c$, if $c \in (-1,\ 0)$. The absence of a periodic component of the trend corresponds to the parameter value $d = 0$, and the absence of a linear component corresponds to – $c = 0$.

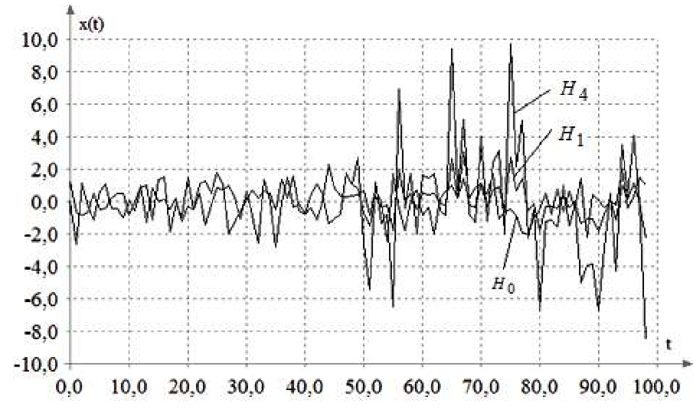

Fig. 7. Trend corresponding to H_0, H_1, H_4 hypotheses

During the analysis of power with respect to linear, periodic, and combined trend in the dispersion characteristics (in variance) of a random variable the following competitive hypotheses were considered:

$$H_5:\ x_i = \xi_i(1 + ct_i),\quad c = 1;$$

$$H_6:\ x_i = \xi_i(1 + d \cdot \sin(2k\pi t_i)),\quad d = 0.8,\quad k = 2;$$

$$H_7: x_i = \xi_i(1 + ct_i + d\sin(2k\pi t_i)),\ c = 1,\ d = 0.8,\ k = 2.$$

Representation of corresponding processes is shown in Figures.◦8 – 10.

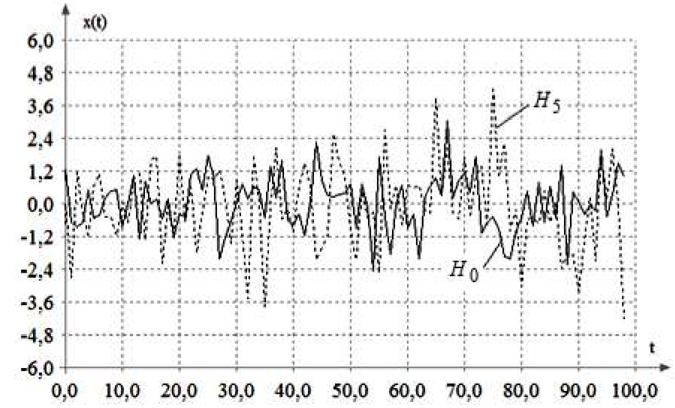

Fig. 8. Linear trend in the dispersion characteristics for H_5

In the course of work statistical simulation methods (for probabilities of errors of the first kind $\alpha = 0.15, 0.1, 0.05, 0.01$) provided estimations of power of the investigated criteria against the competitive hypotheses H_1, H_2, H_3 and H_4 (corresponding to the shift of the dispersion value). Test powers of competing hypotheses H_5, H_6, H_7 were studied. Such hypotheses correspond to linear or nonlinear trend presence in the dispersion characteristics of analyzed processes.

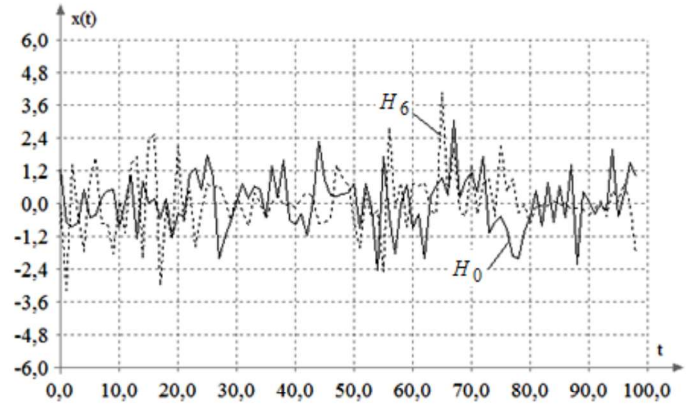

Fig. 9. Periodic trend in the dispersion characteristics for H_6

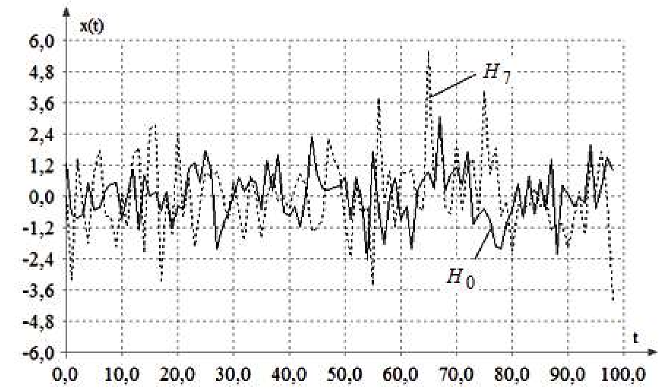

Fig. 10. Combined trend in the dispersion characteristics for H_7

Fig. 11 shows power estimates of Cox-Stewart test against competitive hypotheses H_1, H_2, H_3 and H_4 for value $\alpha = 0.05$.

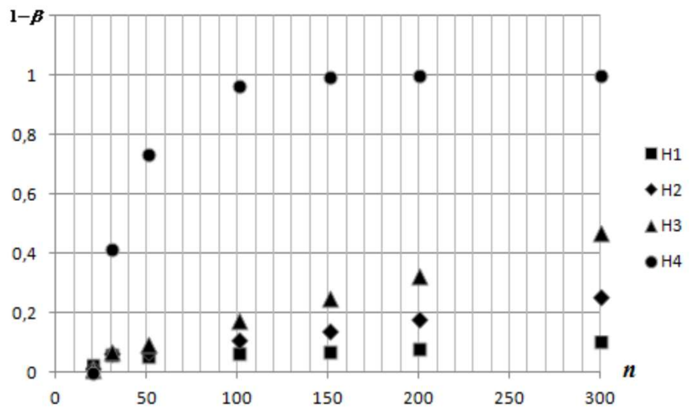

Fig. 11. Power estimates of Cox-Stewart test

Power estimates only with the significance level $\alpha = 0.1$ and the sample number *n=100* are shown in Fig.12 for comparative analysis of power against competitive hypotheses H_1, H_2, H_3 and H_4.

For similar competitive hypotheses criteria Hsu tests with H and G statistics as well as Klotz test showed the highest power with respect to the analyzed sets of competitive hypotheses. They showed the ability to detect trend in the dispersion characteristics when it has a 10% increase.

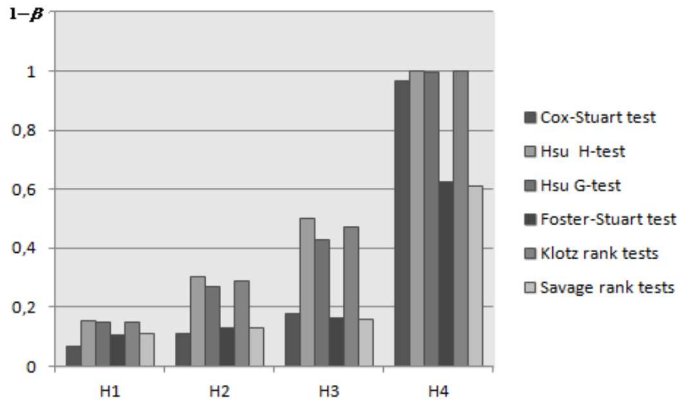

Рис. 11. Power estimates for similar competitive hypotheses

Power estimates only with the significance level $\alpha = 0.1$ and the sample number *n=100* are shown in Table 1 for comparative analysis of power against competitive hypotheses H_5, H_6 and H_7.

Tests are ordered by decreasing power $1-\beta$. Hsu tests with H - and G-statistics and Klotz test can also well distinguish between the null hypothesis and its competitive hypotheses (H_5 and H_6) which exhibit the presence of linear or periodic trend in distribution characteristics.

At the same time Cox–Stuart, Savage and Foster–Stuart tests cannot detect the presence of a periodic trend in the variance reliably (due to relatively low power against similar enough hypothesis H_6).

Unfortunately, none of these tests has shown the ability to detect a combined trend in the dispersion corresponding to the studied hypothesis H_7y. The powers with respect to such hypothesis were extremely low.

TABLE I
COMPARATIVE ANALYSIS OF POWERS OF ALL RANDOMNESS TESTS AND TESTS AGAINST TREND ABSENCE IN VARIANCE

№	**Against** H_5	$1-\beta$	**Against** H_6	$1-\beta$
1	Hsu H	0.836	Hsu H	0.711
2	Hsu G	0.818	Klotz	0.678
3	Klotz	0.807	Hsu G	0.545
4	Cox-Stuart	0.489	Savage	0.196
5	Foster-Stuart	0.346	Cox-Stuart	0.143
6	Savage	0.246	Foster-Stuart	0.048

№	**Against** H_7	$1-\beta$
1	Hsu H	0.162
2	Savage	0.095
3	Foster-Stuart	0.082
4	Hsu G	0.057
5	Cox-Stuart	0.052
6	Klotz	0.104

V. SUMMARY AND CONCLUSION

Thus, methods of statistical modeling have been used to study the statistics distribution of various parametric and nonparametric tests for randomness and the absence of a trend in the dispersion characteristics; within the framework of developing ISW software an interactive study mode of the distributions of the statistics has been implemented for the case of violation of standard assumptions. A comparative analysis of test powers against some competitive hypotheses has been carried out, and results of such analysis can be used to estimate the desirability of application of particular test. Disadvantages of individual criteria have been noted.

Research was conducted with support of the Ministry of Education and Science of the Russian Federation (project #2.541.2014K).

REFERENCES

[1] Kobzar, A.I. Applied mathematical statistics for engineers and academic researchers/ Kobzar A.I.. – M. : Fizmatlit, 2006. – P. 816.

[2] Veretelnikova, I.V., Lemeshko, B.Yu. Analytical review of tests for randomness and trend absence // Materials of the XII international conference "Actual problems of electronic instrument engineering" (APEIE-2014). V.6, Novosibirsk, 2014. – pp. 16-23..

[3] Lemeshko B.Yu., Komissarova A.S., Tsheglov A.E. Application of tests for trend detection and checking for randomness // Metrologiya. 2010. № 12. – pp. 3-25.

[4] Cox D.R., Stuart A. Quick sign tests for trend in location and dispersion // Biometrica. 1955. – V.42. – pp.80-95.

[5] Hsu D. A. Test for variance shift at an unknown time point / D. A. Hsu // Appl. Statist., 1977. – V.26, № 3. – pp. 279–284.

[6] Lemeshko B.Yu., Komissarova A.S., Tsheglov A.E. Properties and power of tests for trend detection and checking for randomness // Scientific Bulletin of NSTU. – 2012. – pp. 53-66.

Veretel`nikova Irina V.,
Postgraduate, Novosibirsk State Technical University, Department of Theoretical and Applied Informatics

Lemeshko Boris Yu.,
Professor, Novosibirsk State Technical University, Department of Theoretical and Applied Informatics