Обработка информации и математическое моделирование : материалы РНТК. Новосибирск, 2016.

УДК 519.24

О мощности модификации критерия проверки равномерности Андерсона-Дарлинга

П.Ю. Блинов, Б.Ю. Лемешко[1]
Новосибирский государственный технический университет

Рассмотрена модификация непараметрическог критерия согласия Андерсона-Дарлинга, предназначенная для проверки равномерности. Исследованы распределения статистики и мощность критерия. Проведен сравнительный анализ мощности критериев. Сделаны рекомендации по применению.

Ключевые слова: равномерное распределение, статистика критерия, мощность критерия.

1. Введение

На проверку равномерности ориентировано множество критериев, обилие которых обусловлено интересом, проявляемым к использованию модели равномерного закона в различных приложениях. Использование такой простой модели во многих ситуациях позволяет найти решение задачи с опорой только на аналитические методы. Если применение модели равномерного закона обосновано, то многие статистические выводы оказываются проще.

Пусть случайные величины $X_1, X_2, ..., X_n$ принадлежат некоторому закону с функцией распределения вероятностей $F(x)$, тогда случайные величины $Y_i = F(X_i)$, $i = \overline{1,n}$ распределены равномерно на интервале $[0,1]$. В связи с этим во многих случаях задачу проверки гипотезы о принадлежности выборки $X_1, X_2, ..., X_n$ некоторому непрерывному закону распределения можно заменить задачей проверки гипотезы о принадлежности выборки $Y_1, Y_2, ..., Y_n$ равномерному закону.

В данной работе методами статистического моделирования исследуются модификация критерия согласия Андерсона-Дарлинга [1], ориентированная на проверку равномерности. Модифицированный критерий сравнивается по мощности с классическим критерие.

При исследовании распределений статистик соответствующих критериев количество экспериментов, осуществляемых при статистическом моделировании принималось равным 1 660 000. Такое количество экспериментов позволяет, с одной стороны, проследить качественную картину, отражающую изменение распределений статистик в зависимости от различных факторов, с другой – обеспечить приемлемую точность получаемых оценок мощности и искомых вероятностей.

В качестве конкурирующего закона при проверке равномерности выбрано бета-распределение 1-го рода $B_I(\theta_0, \theta_1, \theta_2, \theta_3)$ с функцией плотности

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

Так как наибольший интерес вызывает способность критериев различать близкие альтернативы, оценки мощности (как и в [2, 3]) были получены относительно трех конкурирующих

[1] Работа выполнена при поддержке Министерства образования и науки РФ в рамках проектной части государственного задания (№ 2.541.2014/К).

гипотез H_1, H_2, H_3 вида:

$$H_1 : F(x) = B_I(1.5, 1.5, 1, 0), \ x \in [0,1];$$

$$H_2 : F(x) = B_I(0.8, 1, 1, 0), \ x \in [0,1];$$

$$H_3 : F(x) = B_I(1.1, 0.9, 1, 0), \ x \in [0,1].$$

Функции распределения вероятностей, соответствующие рассматриваемым гипотезам, достаточно близки (см. рис. 1), а плотности существенно различаются (см. рис. 2).

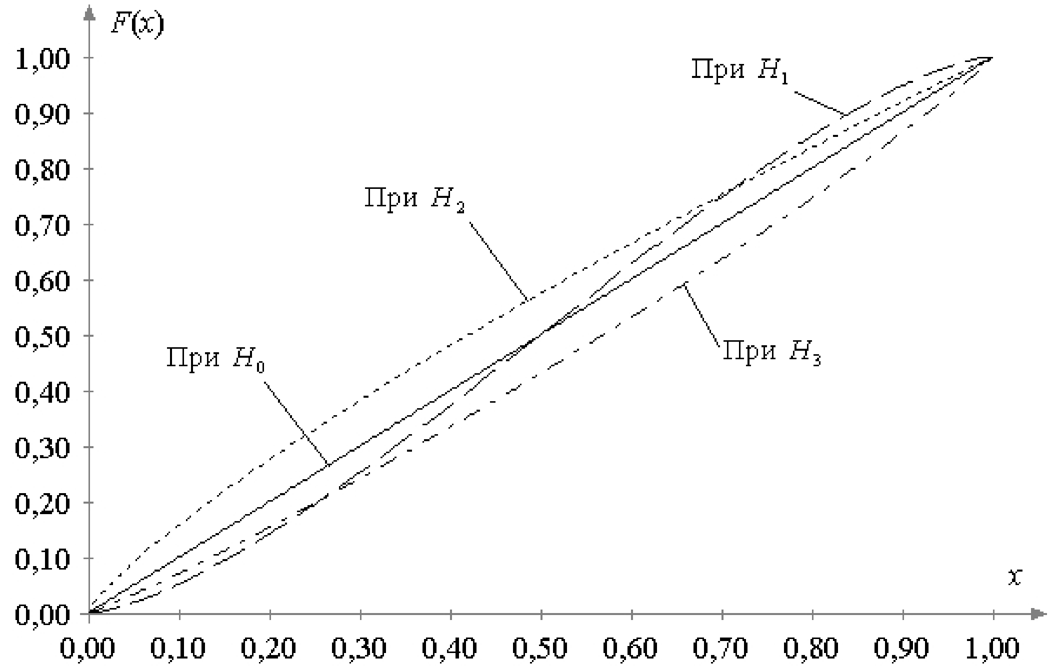

Рис. 1. Функции распределения вероятностей, соответствующие конкурирующим гипотезам

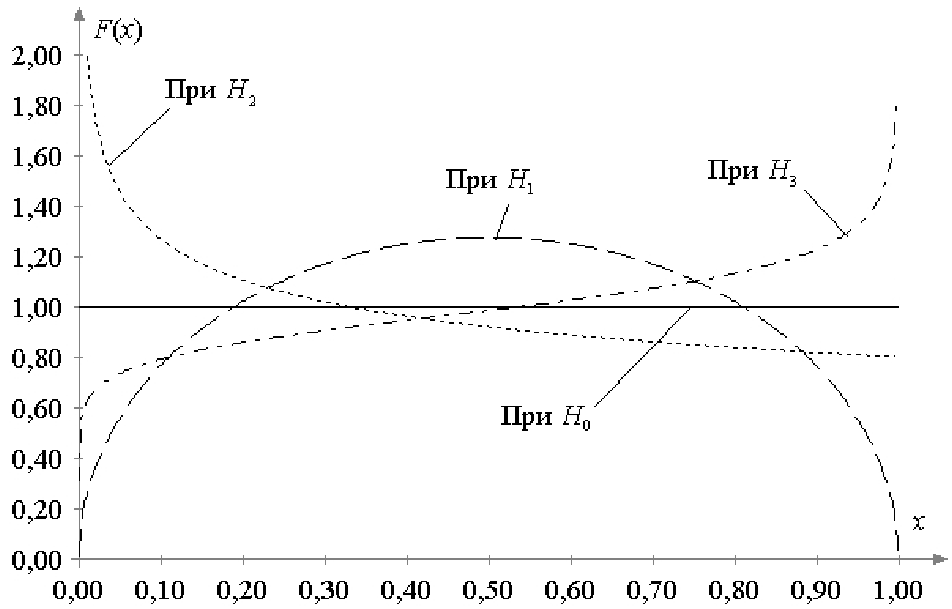

Рис. 2. Плотности распределения законов, соответствующие конкурирующим гипотезам

Конкурирующей гипотезе H_1 соответствует закон, функция распределения которого пересекается с функцией распределения равномерного закона, а при H_2 и H_3 функции распределения законов лежат выше и ниже функции равномерного. Заметим, что при малых объёмах выборок n для многих критериев, в том числе для непаметрических критериев согласия, различать гипотезы H_1 и H_0 представляет значительную трудность. Это связано с обнаруженной в [4] смещенностью соответствующих критериев в этой ситуации (при малых n и малых α).

2. Критерий Андерсона–Дарлинга

Статистика классического критерия согласия Ω^2 Андерсона–Дарлинга [4] при проверке равномерности принимает вид

$$S_\Omega = -n - 2\sum_{i=1}^{n}\left\{\frac{2i-1}{2n}\ln U_i + \left(1-\frac{2i-1}{2n}\right)\ln(1-U_i)\right\}. \qquad (1)$$

При справедливости простой проверяемой гипотезы H_0 (как в данном случае) статистика (1) в пределе подчиняется закону с функцией распределения $a2(s)$, имеющей вид [5]

$$a2(s) = \frac{\sqrt{2\pi}}{s}\sum_{j=0}^{\infty}(-1)^j \frac{\Gamma(j+1/2)(4j+1)}{\Gamma(1/2)\Gamma(j+1)}\exp\left\{-\frac{(4j+1)^2\pi^2}{8s}\right\}\times$$

$$\times\int_0^\infty \exp\left\{\frac{s}{8(y^2+1)} - \frac{(4j+1)^2\pi^2 y^2}{8s}\right\}dy\,. \qquad (2)$$

Зависимостью распределения статистики (1) от объема выборки можно пренебречь при $n \ge 25$. При таких объемах выборок отклонение реального распределения статистики от предельного (2) незначительно и не влияет на результаты статистического вывода.

Проверяемая гипотеза H_0 о равномерности закона не отвергается, если достигнутый уровень значимости $P\left\{S_\Omega > S_\Omega^*\right\} = 1 - a1\left(S_\Omega^*\right)$ не превышает заданной вероятности ошибки 1-го рода α, где S_Ω^* – вычисленное по выборке значение статистики (1). Менее обосновано судить о результатах проверки можно, сравнивая S_Ω^* с критическими значениями из таблицы 1.

Таблица 1. Процентные точки распределения (2) статистики Андерсона-Дарлинга

Функция распределения	Верхние процентные точки				
	0.15	0.1	0.05	0.025	0.01
$a2(S)$	1.6212	1.9330	2.4924	3.0775	3.8781

Оценки мощности критерия Андерсона–Дарлинга при проверке равномерности по отношению к конкурирующей гипотезе H_1 приведены в таблице 2, оценки мощности по отношению к гипотезам H_2 и H_3 – в таблицах 3 и 4 соответственно.

В общем случае непараметрический критерий Андерсона–Дарлинга представляет собой один из наиболее мощных критериев согласия [6, 7]. В этом можно убедиться в том числе, сравнивая оценки его мощности относительно гипотез H_2 и H_3 с оценками других непараметрических критериев согласия. Однако в данном случае по отношению к рассматриваемой гипотезе H_1 следует отметить существенную смещённость критерия, о чём можно судить по оценкам мощности при малых n (см. табл. 2, тёмным цветом выделены ячейки таблицы с оценками мощности $1-\beta \le \alpha$).

Таблица 2. Мощность критерия Андерсона-Дарлинга относительно гипотезы H_1

n	α				
	0.15	0.1	0.05	0.025	0.01
10	0.095	0.053	0.019	0.007	0.002
20	0.140	0.078	0.028	0.010	0.003
30	0.196	0.114	0.042	0.014	0.004
40	0.258	0.156	0.060	0.021	0.005
50	0.325	0.206	0.084	0.031	0.007
100	0.652	0.505	0.283	0.134	0.041
150	0.861	0.760	0.544	0.332	0.138
200	0.954	0.904	0.762	0.565	0.311
300	0.998	0.990	0.959	0.882	0.702

Таблица 3. Мощность критерия Андерсона-Дарлинга относительно гипотезы H_2

n	α				
	0.15	0.1	0.05	0.025	0.01
10	0.256	0.192	0.116	0.071	0.036
20	0.324	0.253	0.165	0.107	0.059
30	0.390	0.315	0.215	0.145	0.084
40	0.449	0.371	0.263	0.184	0.111
50	0.505	0.426	0.313	0.225	0.141
100	0.718	0.648	0.533	0.426	0.308
150	0.846	0.795	0.700	0.601	0.475
200	0.919	0.885	0.817	0.737	0.624
300	0.980	0.968	0.938	0.897	0.829

Таблица 4. Мощность критерия Андерсона-Дарлинга относительно гипотезы H_3

n	α				
	0.15	0.1	0.05	0.025	0.01
10	0.209	0.151	0.086	0.049	0.023
20	0.262	0.197	0.119	0.072	0.037
30	0.315	0.244	0.157	0.099	0.053
40	0.363	0.289	0.192	0.126	0.069
50	0.410	0.333	0.230	0.154	0.088
100	0.604	0.526	0.406	0.303	0.200
150	0.742	0.675	0.560	0.451	0.325
200	0.836	0.783	0.685	0.582	0.452
300	0.937	0.909	0.849	0.776	0.668

На рис. 3 показаны распределение $G(S|H_0)$ статистики критерия Андерсона–Дарлинга при справедливости проверяемой гипотезы H_0 и распределения $G(S_n|H_1)$ этой статистики при справедливости H_1 (для объемов выборок $n = 10, 20, 100, 300$). Как можно видеть, распределения статистики $G(S_n|H_1)$ при $n = 10, 20$ пересекают $G(S|H_0)$, что объясняет почему мощность $1-\beta$ оказывается меньше α

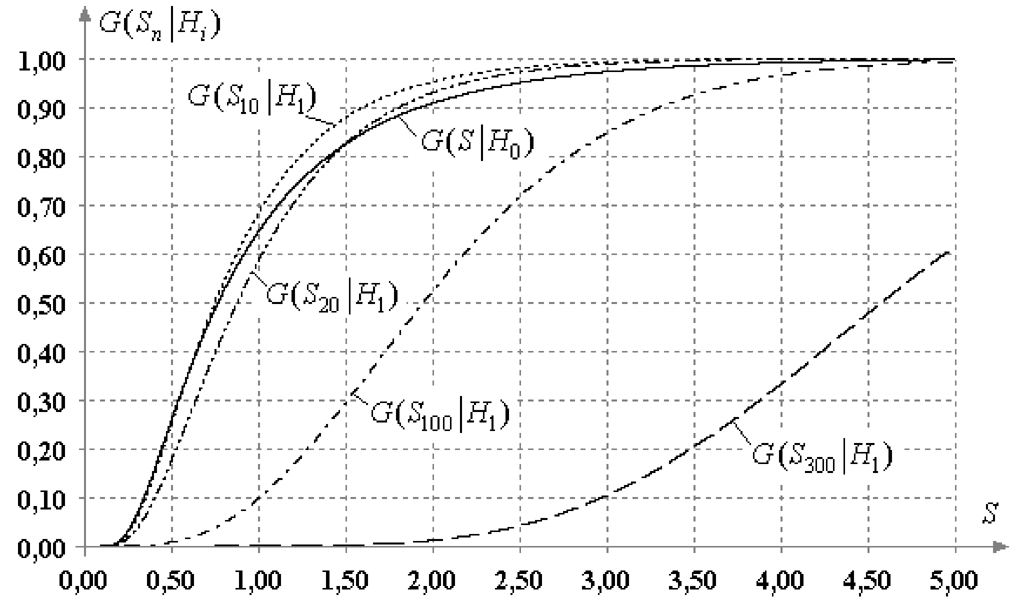

Рис. 3. Распределения $G(S|H_0)$ и $G(S_n|H_1)$ статистики критерия Андерсона–Дарлинга при проверке равномерности

На рисунке распределение $G(S|H_0)$ показано только при $n=10$. При $n \ge 20$ распределения $G(S_n|H_0)$ визуально не отличаются от приведенного на рисунке $G(S_{10}|H_0)$ и практически совпадают с предельным распределением $a2(s)$ статистики критерия согласия Андерсона–Дарлинга при проверке простых гипотез.

3. Модифицированный критерий Андерсона-Дарлинга

Статистику критерия Андерсона-Дарлинга можно представить в интегральном виде

$$W^2 = n\int_{-\infty}^{\infty} \frac{\left[F_n(x) - F(x)\right]^2}{F(x)\left[1-F(x)\right]} dF(x). \tag{3}$$

При проверке равномерности статистику (3) можно записать как

$$W^2 = n\int_{-\infty}^{\infty} \frac{\left[F_n(x) - x\right]^2}{x\left[1-x\right]} dx, \tag{4}$$

заменив $F(x)$ на x и $dF(x)$ на dx. В [1] окончательное выражение для статистики спринимает вид:

$$V^2 = n\left[\frac{(U_1-c_1)^2}{(1-c_1)} + \sum_{i=1}^{n} \frac{(U_i-c_i)^2}{c_i(1-c_i)}(c_{i+1}-c_i) + \frac{(U_n-c_n)^2}{c_n}\right], \tag{5}$$

где $c_i = (i-0.375)/(n+0.25)$.

В таблице 5 представлены уточнённые процентные точки для модификации (5), полученные в процессе моделирования 1 660 000 значений статистики при справедливости гипотезы о равномерности анализируемых выборок. Процентные точки приведенные в таблице 5 несколько отличаются от результатов авторов [1], которые получены при моделировании с меньшей точностью при числе экспериментов метода Монте–Карло 100 000 выборок. (отличие в 3-м знаке после запятой).

Таблица 5. Критические значения статистики (5)

n	$1-\alpha$				
	0.8	0.85	0.9	0.95	0.99
10	1.413	1.676	2.072	2.820	4.884
20	1.467	1.719	2.098	2.796	4.655
30	1.472	1.716	2.080	2.745	4.473
40	1.471	1.713	2.068	2.708	4.358
50	1.467	1.702	2.050	2.679	4.271
100	1.453	1.677	2.008	2.598	4.085
150	1.441	1.662	1.986	2.564	4.010
200	1.435	1.653	1.972	2.548	3.978
300	1.428	1.644	1.962	2.532	3.939
1000	1.414	1.628	1.943	2.507	3.900

Наши исследования показали, что распределение модифицированной статистики (5) сходится к предельному распределению $a2(s)$ значительно медленнее. В частности, гипотеза о согласии эмпирического распределения (при числе экспериментов $N=10^6$) с теоретическим распределением $a2(s)$ не отвергается (при заданном уровне значимости $\alpha = 0.05$) при объёмах выборок $n > 1000$. При числе экспериментов $N = 1.66\times10^4$ – гипотеза о согласии не отвергается при $n \geq 300$.

Оценки мощности критерия со статистикой (5) относительно конкурирующей гипотезы H_1 представлены в таблице 6, относительно гипотез H_2 и H_3 – в таблицах 7 и 8 соответсвенно.

Таблица 6. Мощность критерия относительно гипотезы H_1

n	α				
	0.15	0.1	0.05	0.025	0.01
10	0.199	0.131	0.062	0.029	0.010
20	0.297	0.208	0.109	0.055	0.022
30	0.386	0.283	0.159	0.085	0.036
40	0.467	0.356	0.212	0.119	0.053
50	0.544	0.429	0.268	0.157	0.073
100	0.820	0.730	0.560	0.396	0.227
150	0.943	0.897	0.784	0.638	0.441
200	0.985	0.967	0.911	0.818	0.649
300	0.999	0.998	0.990	0.969	0.908

Стоит заметить, что при в ситуации проверки равномерности с конкурирующей гипотезой H_1 смещенность не проявляется, в отличие от непараметрического критерия Андерсона-Дарлинга.

4. Сравнительный анализ мощности критериев

Мощность модифицированного критерия Андерсона–Дарлинга сравнивалась с мощность разлчных критериев, используемых при проверке равномерности.

Таблица 7. Мощность критерия относительно гипотезы H_2

n	α

	0.15	0.1	0.05	0.025	0.01
10	0.195	0.135	0.070	0.035	0.013
20	0.257	0.189	0.110	0.062	0.028
30	0.319	0.245	0.153	0.093	0.047
40	0.378	0.299	0.197	0.126	0.068
50	0.435	0.354	0.243	0.162	0.091
100	0.662	0.585	0.461	0.352	0.236
150	0.810	0.750	0.642	0.531	0.399
200	0.897	0.856	0.773	0.680	0.551
300	0.972	0.957	0.919	0.866	0.781

Таблица 8. Мощность критерия относительно гипотезы H_3

n	α				
	0.15	0.1	0.05	0.025	0.01
10	0.216	0.157	0.092	0.054	0.026
20	0.263	0.198	0.120	0.072	0.036
30	0.312	0.240	0.152	0.094	0.049
40	0.358	0.283	0.186	0.120	0.065
50	0.404	0.326	0.221	0.147	0.083
100	0.598	0.519	0.398	0.295	0.191
150	0.738	0.669	0.553	0.442	0.317
200	0.833	0.779	0.678	0.574	0.442
300	0.935	0.907	0.846	0.771	0.661

В таблице 9 по величине мощности, показанной относительно конкурирующих гипотез H_1, H_2, H_3, упорядочены критерии, показавшие в [2, 4] удовлетворительную мощность, позволяющую рекомендовать их для применения. Кроме рассмотренных в данной работе к данному подмножеству критериев проверки равномерности относятся [2, 4] следующие критерии: Жанга, Неймана-Бартона, χ^2 Пирсона, Фросини, Крамера–Мизеса–Смирнова, Шермана, энтропийный критерий Дудевича-ван дер Мюлена и его модификации.

В таблице рассмотренные критерии упорядочены по убыванию мощности относительно соответствующих конкурирующих гипотез (по величине мощности $1-\beta$, проявленной при $n=100$ и уровне значимости $\alpha=0.1$). В столбце для H_1 темным цветом выделены критерии, которые относительно H_1 при малых n обладают ярко выраженной смещённостью.

Таблица 9. Упорядоченность критериев по мощности

№ п/п	Относительно H_1	$1-\beta$	Относительно H_2	$1-\beta$	Относительно H_3	$1-\beta$
1	Модификация энтропийного критерия 2	0.883	Андерсона–Дарлинга	0.648	Андерсона–Дарлинга	0.526
2	Жанга Z_A	0.850	Жанга Z_C	0.606	Фросини	0.522
3	Неймана–Бартона N_2	0.837	Фросини	0.603	***Модификация Андерсона-Дарлинга***	***0.519***

Окончание таблицы 9.

№ п/п	Относительно H_1	$1-\beta$	Относительно H_2	$1-\beta$	Относительно H_3	$1-\beta$
4	Жанга Z_C	0.819	Неймана–Бартона N_2	0.597	Крамера–Мизеса–Смирнова	0.507
5	Дудевича–ван дер Мюлена	0.790	Крамера–Мизеса–Смирнова	0.595	Жанга Z_C	0.463
6	Модификация энтропийного критерия 1	0.789	Жанга Z_K	0.590	Жанга Z_A	0.459
7	***Модификация Андерсона-Дарлинга***	***0.730***	***Модификация Андерсона-Дарлинга***	***0.585***	Неймана–Бартона N_2	0.447
8	Жанга Z_K	0.617	Жанга Z_A	0.574	Жанга Z_K	0.438
9	χ^2 Пирсона	0.593	χ^2 Пирсона	0.448	χ^2 Пирсона	0.374
10	Андерсона–Дарлинга	0.505	Модификация энтропийного критерия 1	0.328	Дудевича–ван дер Мюлена	0.275
11	Фросини	0.384	Дудевича–ван дер Мюлена	0.327	Модификация энтропийного критерия 1	0.275
12	Крамера–Мизеса–Смирнова	0.358	Модификация энтропийного критерия 2	0.266	Модификация энтропийного критерия 2	0.267
13	Шермана	0.215	Шермана	0.204	Шермана	0.154

Таким образом, можно отметить, что модификация критерия Андерсона-Дарлинга, предложенная в [1] для проверки равномерности, является весьма эффектиным критерием. В качестве недостатка критерия следует указать медленную сходимость распределенич статистики к предельному закону. А в качестве преимущества – отсутствия смещённости относительно конкурирующей гипотезы H_1.

Литература

1. Rahman Mezbahur, Larry M. Pearson, Herbert C Heien. A Modifed Anderson-Darling Test for Uniformity// Bulletin of the Malaysian Mathematical Sciences Society. Second Series, Vol 29, P. 11-16.
2. *Блинов П.Ю., Лемешко Б.Ю.* О критериях проверки отклонения распределения от равномерного закона //Материалы Российской НТК “Обработка информации и математическое моделирование”, Новосибирск. 2015. – С. 21-31.
3. *Blinov P.Yu., Lemeshko B.Yu.* The Comparative Analysis of Tests in the Problem of Testing the Hypothesis of Uniformity // Proceedings of the International Workshop “Applied Methods of Statistical Analysis. Nonparametric Approach” – AMSA’2015, Novosibirsk–Belokuricha, Russia, 14-19 September, 2015. – P. 92-100.
4. *Лемешко Б.Ю., Блинов П.Ю.* Критерии проверки отклонения распределения от равномерного закона. Руководство по применению. – М.: ИНФРА-М, 2015. – 183 с. – (Научная мысль). DOI: 10.12737/11304
5. *Большев Л. Н. Смирнов Н. В.* Таблицы математической статистики. – М. : Наука, 1983. – 416 с.
6. *Лемешко Б. Ю.* Непараметрические критерии согласия. Руководство по применению. – М.: ИНФРА-М, 2014. – 163 с. DOI: 10.12737/11873
7. *Лемешко Б. Ю.* Критерии проверки отклонения распределения от нормального закона: Руководство по применению. – М.: ИНФРА-М, 2015. – 160 с. – (Научная мысль). DOI: 10.12737/6086.

Статья поступила в редакцию 28.02.2015;

Лемешко Борис Юрьевич

Д.т.н., профессор, г.н.с. кафедры теоретической и прикладной информатики НГТУ (630073, Новосибирск, просп. Карла Маркса, 20), e-mail: lemeshko@ami.nstu.ru

Блинов Павел Юрьевич

Аспирант кафедры теоретической и прикладной информатики НГТУ (630073, Новосибирск, просп. Карла Маркса, 20), e-mail: blindizer@ya.ru.

The power of modified Anderson-Darling test for uniformity testing

P. Yu. Blinov, B. Yu. Lemeshko

Novosibirsk state Technical university

The modification of non-parametric Anderson-Darling test intended for testing of uniformity have been considered. Distributions of test statistics, the power of tests have been studied. The comparative analysis of test power have been carried out. The recommendations on the application have been made.

Key words: uniform distibbution, testing hypothesis, test staticstic, power of test