

Сибирский государственный университет  
телекоммуникаций и информатики

**ОБРАБОТКА ИНФОРМАЦИИ  
И  
МАТЕМАТИЧЕСКОЕ  
МОДЕЛИРОВАНИЕ**

РОССИЙСКАЯ  
НАУЧНО-ТЕХНИЧЕСКАЯ  
КОНФЕРЕНЦИЯ

МАТЕРИАЛЫ КОНФЕРЕНЦИИ

Новосибирск  
2018

**ISBN 978-5-91434-042-8**

© ФГБОУ ВО «Сибирский государственный университет телекоммуникаций и информатики» 2018  
© Авторы 2018

# СОДЕРЖАНИЕ

## Секция 1

### ИНФОРМАТИКА И МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

#### Подсекция 1.1. НГТУ

<b>Бауэр Д.В., Гультяева Т.А.</b> Разработка программной системы электронного голосования на децентрализованной платформе.	6
<b>Блинов П.Ю., Лемешко Б.Ю.</b> Свойства критериев экспоненциальности Дешпанде.	10
<b>Гриф А.М.</b> Экологический 3D-мониторинг качества воздуха города Новосибирска на основе данных спутниковой навигации, мобильных экометрических станций и метода конечных элементов.	17
<b>Зорина А.А., Лемешко Б.Ю.</b> О критериях проверки показательности Аткинсона.	26
<b>Кобьялянский В.Г., Михед К.А.</b> Исследование динамических характеристик виртуального прибора ColorLearn среды LabVIEW.	31
<b>Кочнев А.В., Волкова В.М.</b> Идентификация сообществ, формируемых системой горизонтального премиривания методами кластеризации в графах.	35
<b>Лемешко Б.Ю., Белоцерковец В.Н.</b> О свойствах и мощности критериев нормальности Лина–Мудхолкара и Васичека.	40
<b>Лемешко Б.Ю., Веретельникова И.В.</b> О применении и мощности $k$ -выборочных критериев однородности законов.	48
<b>Лемешко Б.Ю., Новикова А.Ю.</b> О критериях Миллера и Лайарда и мощности критериев однородности дисперсий.	60
<b>Морозов Ю.В., Спектор А.А.</b> Выравнивание амплитуд импульсов шагов человека при классификации сейсмических сигналов.	70
<b>Осинцева Е.А., Чимитова Е.В.</b> Построение оптимальных планов эксперимента на основе винеровской деградационной модели.	75
<b>Патрушев И.И., Персова М.Г., Соловейчик Ю.Г.</b> Исследование численного метода трёхмерного моделирования процесса многофазной фильтрации.	85
<b>Поверин Д.В., Постовалов С.Н.</b> Оценивание вероятности обнаружения новых ассоциаций при комбинировании результатов полногеномного анализа ассоциаций.	93
<b>Попов А.А., Бобоев Ш.А.</b> Сравнение разреженных решений, получаемых разбиением выборки на части на основе внешних критериев качества моделей в методе LS–SVM.	102
<b>Попов А.А., Холдонов А.А.</b> Построение деревьев регрессии при разбиении области действия факторов на нечеткие партиции.	110
<b>Попов А.А., Холкин В.В.</b> Построение робастных и разреженных решений по методу опорных векторов с функцией потерь Йохана Сайкинса.	117
<b>Сергеева С.А., Чимитова Е.В.</b> Построение обратной гауссовской деградационной модели с фиксированным и случайным эффектами.	123
<b>Соснин И.В., Гультяева Т.А.</b> Применение NLP-библиотек для решения задач классификации текстов.	135
<b>Толстобров И.А., Ступаков И.М.</b> Вычисление сингулярных интегралов для базисных функций высокого порядка в методе граничных элементов с применением рекуррентных соотношений.	139
<b>Филоненко П.А., Постовалов С.Н.</b> Выбор статистического критерия однородности распределений с помощью правила Сэвиджа для принятия решений в условиях риска и неопределенности.	144
<b>Черникова О.С., Долгов А.А.</b> Применение адаптивного сигма-точечного фильтра Калмана при исследовании непрерывно-дискретных систем.	150
<b>Чубич В.М., Прокофьева А.Э.</b> Активная параметрическая идентификация одной динамической системы с использованием робастного оценивания.	159

# Свойства критериев экспоненциальности Дешпанде

П. Ю. Блинов, Б. Ю. Лемешко<sup>1</sup>

Новосибирский государственный технический университет

Исследуются свойства критерия Дешпанде, предназначенного для проверки гипотез о принадлежности наблюдений экспоненциальному закону. Исследуются распределения статистик и мощность критерия относительно различных конкурирующих гипотез в зависимости от выбора параметра  $b$ . Мощность критерия сравнивается с мощностью других критериев экспоненциальности. Даются рекомендации о применении критерия.

*Ключевые слова:* экспоненциальный закон, проверка гипотез, мощность критерия

## 1. Введение

Экспоненциальный закон распределения вероятностей является одной из базовых моделей, используемых в теории надежности. Его аналитическая простота делает его привлекательным для инженеров и исследователей. Однако применение любой модели требует обоснования. И в данном случае, прежде чем использовать экспоненциальный закон в качестве модели надежности следует убедиться в том, что наблюдаемая случайная величина (например, моменты отказов) хорошо описывается этим законом. В ином случае выигрыш от простоты расчетов будет «компенсирован» некорректностью выводов статистического анализа, вызванной отклонением реального закона распределения от экспоненциального.

Для проверки гипотезы о принадлежности выборки показательному закону распределения предложено множество статистические критерии. Обилие критериев обусловлено, с одной стороны, популярностью использования этой модели в приложениях, а с другой, отсутствием критерия экспоненциальности, обладающего наибольшей мощностью относительно любой альтернативы (относительно любой конкурирующей гипотезы).

В этой связи актуальным вопросом оказывается знание детальных свойств критериев (распределений статистик, мощности относительно близких конкурирующих гипотез) и возможность ранжирования критериев, чтобы в конкретной ситуации выбрать наиболее предпочтительный, который с заданной вероятностью ошибки первого рода минимизировал бы вероятность ошибки второго рода.

В данной работе исследуются свойства и мощность семейства критериев Дешпанде [1], которые сравниваются по мощности с рядом других популярных критериев экспоненциальности. В процессе исследований методом Монте–Карло распределений статистик и мощности критериев количество имитационных экспериментов принималось равным 1 660 000.

## 2. Описание критерия

Пусть  $x_1, x_2, \dots, x_n$  – выборка н.о.р. случайных величин. Проверяется гипотеза  $H_0$  о принадлежности её показательному закону с плотностью  $f(x) = \exp(-x)$ .

<sup>1</sup> Исследования выполнены при поддержке Министерства образования и науки РФ в рамках государственной работы «Обеспечение проведения научных исследований» (№ 1.4574.2017/6.7) и проектной части государственного задания (№ 1.1009.2017/4.6).

Дешпанде [1] было предложено семейство критериев показательности, где в качестве альтернативы рассматривались законы распределения с возрастающей интенсивностью отказов. Статистики критериев Дешпанде описывается формулой:

$$J = \frac{1}{n(n-1)} \sum_{i \neq j} \varphi(x_i; bx_j), \quad (1)$$

где

$$\varphi(x_i, bx_j) = \begin{cases} 1 & \text{если } x_i > bx_j; \\ 0 & \text{если } x_i \leq bx_j. \end{cases}$$

Критерий является двусторонним, гипотеза о показательности отвергается как при больших, так и при малых значениях статистики (1). Дискретность распределения статистики хорошо заметна при объёмах выборки  $n = 20$ , однако становится несущественной при  $n \geq 50$ .

Для критерия автором предложена нормализация статистики в виде:

$$\sqrt{n} \left( J - \frac{1}{b+1} \right) / \left( 1 + \frac{b}{b+2} + \frac{1}{2b+1} + \frac{2(1-b)}{b+1} - \frac{2b}{b^2+b+1} - \frac{4}{(b+1)^2} \right), \quad (2)$$

которая достаточно хорошо аппроксимируется стандартным нормальным законом.

На рисунках 1 и 2 представлены распределения статистики (2) критерия Дешпанде для  $b = 0.1$  при объёмах выборок 50 и 100 и стандартное нормальное распределение. Распределения довольно близки. Однако следует заметить, что сходимость к нормальному закону зависит от параметра  $b$ .

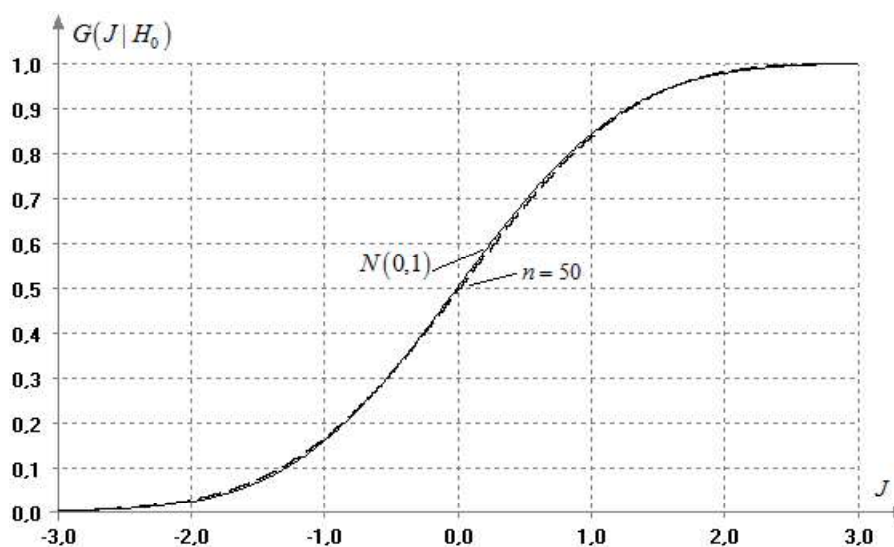


Рис. 1. Распределение статистики (2) при  $n = 50$  и стандартный нормальный закон

В различных работах предлагаются различные оптимальные значения параметра  $b$ . Так в [2] предлагается использовать значение 0.44, которое предпочтительнее против альтернативы, соответствующей распределению Вейбулла. Автор критерия [1] предлагает использовать значения 0.5 или 0.9. Однако наше исследование указывает на эффективность применения других значений параметра  $b$  (см. п.4)

### 3. Конкурирующие гипотезы

Показательному закону соответствует постоянная интенсивность отказов. В связи с этим в качестве конкурирующих гипотез при анализе мощности рассматривались законы, принад-

лежащие к трем классам: с возрастающими, убывающими и немонотонными интенсивностями отказов [3].

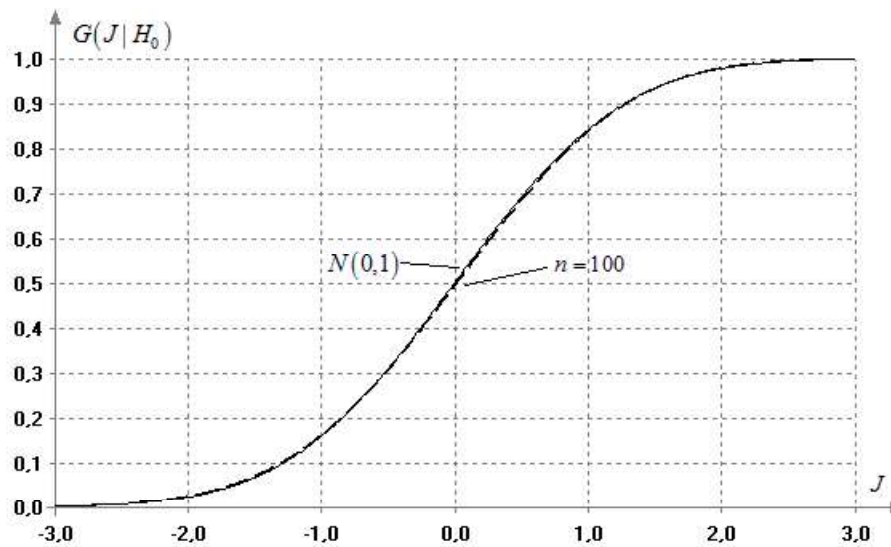


Рис. 2. Распределение статистики (2) при  $n = 100$  и стандартный нормальный закон

В данном случае исследования были проведены для трех конкурирующих гипотез:

- $H_1$  : логнормальное распределение с плотностью  $f(x) = (\theta x \sqrt{2\pi})^{-1} \exp(-(\ln x)^2 / 2\theta^2)$  и с параметром формы 1 (в качестве закона с немонотонной интенсивностью отказов);
- $H_2$  : распределение Вейбулла с плотностью  $f(x) = \theta x^{\theta-1} \exp(-x^\theta)$  и параметром формы 0.7 (в качестве закона с убывающей интенсивностью отказов);
- $H_3$  : распределение Вейбулла с параметром формы 1.2 (в качестве закона с возрастающей интенсивностью отказов).

Функции распределений, соответствующие конкурирующим гипотезам, приведены на рис. 3.

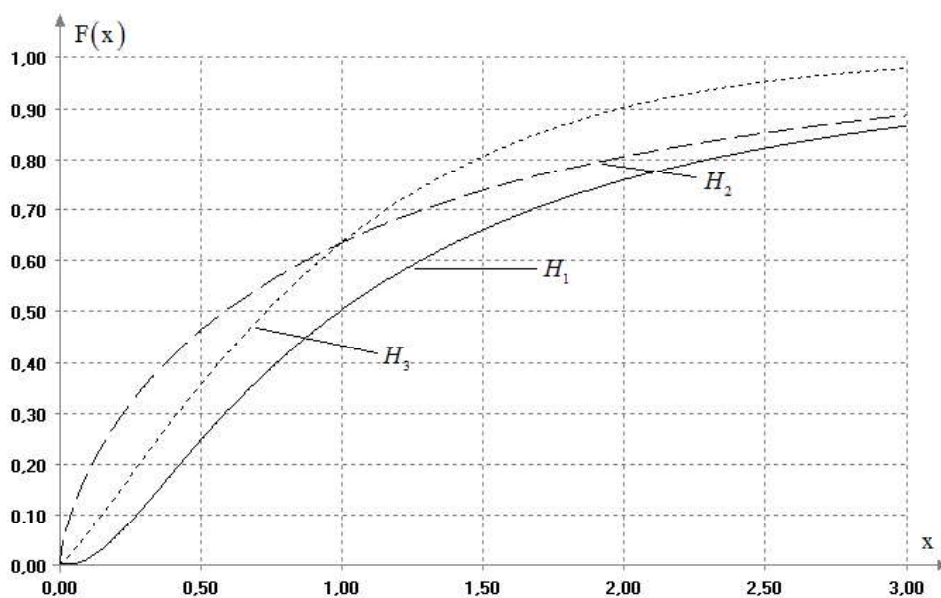


Рис. 3. Распределения, соответствующие конкурирующим гипотезам

#### 4. Анализ мощностей

От выбора параметра  $b$ , используемого в статистике критерия, зависит мощность критерия относительно различных альтернатив.

Для выбора оптимального параметра было проведено исследование мощности при размере входных выборок  $n = 50$ . Полученные оценки мощности представлены в таблице 1. Максимальные мощности в столбце (относительно соответствующей гипотезы) выделены жирным шрифтом.

Таблица 1. Оценки мощности критерия Дешпанде при различных значениях параметра  $b$

$J(b)$	$H_1$			$H_2$			$H_3$		
	$\alpha$			$\alpha$			$\alpha$		
	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
$J(0.1)$	<b>0.482</b>	<b>0.329</b>	<b>0.110</b>	0.874	0.807	0.621	0.355	0.241	0.089
$J(0.2)$	0.371	0.243	0.080	<b>0.891</b>	0.829	0.663	0.398	0.275	0.106
$J(0.3)$	0.309	0.194	0.059	0.890	<b>0.831</b>	<b>0.669</b>	<b>0.408</b>	<b>0.284</b>	<b>0.109</b>
$J(0.4)$	0.265	0.164	0.048	0.888	0.830	0.659	0.402	0.280	0.106
$J(0.5)$	0.243	0.144	0.040	0.881	0.816	0.646	0.399	0.272	0.100
$J(0.6)$	0.224	0.132	0.035	0.865	0.798	0.624	0.387	0.264	0.096
$J(0.7)$	0.210	0.123	0.031	0.835	0.767	0.579	0.371	0.252	0.086
$J(0.8)$	0.193	0.114	0.027	0.788	0.709	0.481	0.341	0.231	0.075
$J(0.9)$	0.170	0.094	0.024	0.641	0.552	0.310	0.287	0.179	0.058

В дальнейших исследованиях мощности критерия рассматривались лишь статистики с двумя значениями параметров  $b = 0.1$  и  $b = 0.3$ , так как именно при этих значениях достигалась максимальная мощность критерия относительно альтернатив  $H_1$  и  $H_2 - H_3$ , соответственно.

В таблицах 2-4 приведены полученные оценки мощности относительно рассматриваемых конкурирующих гипотез для различных объемов выборок от  $n = 10$  до  $n = 300$ .

Таблица 2. Оценки мощности критерия Дешпанде относительно гипотезы  $H_1$

$n$	Дешпанде (0.1)					Дешпанде (0.3)				
	$\alpha$					$\alpha$				
	0.15	0.1	0.05	0.025	0.01	0.15	0.1	0.05	0.025	0.01
10	0.170	0.166	0.166	0.165	0.165	0.156	0.120	0.066	0.037	0.012
20	0.286	0.221	0.111	0.068	0.022	0.221	0.159	0.091	0.048	0.021
30	0.402	0.306	0.187	0.107	0.043	0.284	0.214	0.122	0.071	0.034
40	0.496	0.395	0.251	0.151	0.072	0.341	0.260	0.160	0.094	0.045
50	0.589	0.482	0.329	0.216	0.110	0.394	0.309	0.194	0.118	0.059
100	0.864	0.798	0.669	0.531	0.364	0.621	0.527	0.386	0.271	0.160
150	0.961	0.934	0.866	0.771	0.625	0.777	0.699	0.563	0.434	0.290
200	0.990	0.981	0.953	0.906	0.813	0.873	0.817	0.704	0.584	0.432
300	0.999	0.999	0.996	0.988	0.967	0.963	0.939	0.880	0.802	0.677

Мощность критериев Дешпанде сравнивалась с мощностью других критериев показателности относительно тех же конкурирующих гипотез. Описание этих критериев можно найти в [4, 5, 6].

**Таблица 3. Оценки мощности критерия Дешпанде относительно гипотезы  $H_2$**

$n$	Дешпанде (0.1)					Дешпанде (0.3)				
	$\alpha$					$\alpha$				
	0.15	0.1	0.05	0.025	0.01	0.15	0.1	0.05	0.025	0.01
10	0.430	0.310	0.256	0.158	0.119	0.362	0.283	0.205	0.139	0.086
20	0.617	0.552	0.440	0.331	0.232	0.643	0.564	0.449	0.362	0.250
30	0.757	0.697	0.598	0.501	0.380	0.775	0.719	0.627	0.523	0.412
40	0.851	0.801	0.717	0.626	0.510	0.866	0.824	0.744	0.662	0.559
50	0.907	0.874	0.807	0.731	0.621	0.921	0.890	0.831	0.766	0.669
100	0.993	0.988	0.976	0.958	0.924	0.995	0.992	0.983	0.970	0.944
150	0.999	0.999	0.998	0.995	0.988	1	0.999	0.999	0.997	0.993
200	1	1	1	0.999	0.999	1	1	1	1	0.999
300	1	1	1	1	1	1	1	1	1	1

**Таблица 4. Оценки мощности критерия Дешпанде относительно гипотезы  $H_3$**

$n$	Дешпанде (0.1)					Дешпанде (0.3)				
	$\alpha$					$\alpha$				
	0.15	0.1	0.05	0.025	0.01	0.15	0.1	0.05	0.025	0.01
10	0.211	0.200	0.198	0.195	0.194	0.195	0.154	0.089	0.053	0.019
20	0.264	0.213	0.124	0.084	0.034	0.282	0.213	0.131	0.076	0.036
30	0.329	0.254	0.163	0.103	0.051	0.362	0.286	0.179	0.113	0.060
40	0.382	0.303	0.198	0.126	0.067	0.431	0.347	0.234	0.151	0.082
50	0.442	0.355	0.241	0.162	0.089	0.494	0.408	0.284	0.190	0.109
100	0.664	0.579	0.449	0.335	0.218	0.733	0.656	0.527	0.407	0.276
150	0.805	0.740	0.622	0.504	0.366	0.869	0.815	0.712	0.601	0.458
200	0.891	0.845	0.753	0.651	0.514	0.938	0.905	0.835	0.750	0.624
300	0.968	0.949	0.904	0.843	0.745	0.987	0.978	0.953	0.915	0.844

Приводимые в таблице 5 оценки мощности сравниваемых критериев получены при объеме выборки  $n = 50$ . Оценки мощности находились по эмпирическим распределениям статистик критериев, соответствующих проверяемой и конкурирующим гипотезам. Для получения приемлемой точности эмпирические распределения статистик строились по 1 660 000 испытаниям.

В таблице 5 сравниваемые специальные критерии проверки экспоненциальности упорядочены по убыванию мощности относительно соответствующих конкурирующих гипотез (по величине мощности  $1 - \beta$  при заданном уровне значимости  $\alpha = 0.05$ ).

Относительно гипотезы  $H_1$  критерий Дешпанде уступает критерию Кимбера-Мичела, который показывает наилучший результат, значительно опережая по мощности другие критерии, а также уступает некоторым другим критериям. В то же время в данном случае критерий Дешпанде (0.1) показывает довольно неплохую мощность, а критерий Дешпанде (0.3) опережает критерий Бартлетта–Морана.

Относительно гипотезы  $H_2$  критерии Дешпанде уступают только лидеру критерию Бартлетта–Морана, а относительно гипотезы  $H_3$  критерий Дешпанде (0.3) уступает лидеру в



данной ситуации, критерию Бартлетта–Морана и критерию Фросини, и опережает критерий Кимбера–Мичела.

**Таблица 5. Оценки мощности критериев показательности относительно конкурирующих гипотез  $H_1$ ,  $H_2$  и  $H_3$  при  $n = 50$**

№ п/п	Относительно $H_1$	$1 - \beta$	Относительно $H_2$	$1 - \beta$	Относительно $H_3$	$1 - \beta$
1	Кимбера–Мичела	0.516	Бартлетта–Морана	0.890	Бартлетта–Морана	0.315
2	Корреляционный критерий 2	0.377	<b>Дешпанде (0.3)</b>	0.831	Фросини	0.291
3	Шапиро–Уилка 1	0.359	<b>Дешпанде (0.1)</b>	0.807	<b>Дешпанде (0.3)</b>	0.284
4	Корреляционный критерий 1	0.344	Шермана	0.804	Кимбера–Мичела	0.279
5	<b>Дешпанде (0.1)</b>	0.329	Фросини	0.804	Шермана	0.277
6	Фросини	0.310	Кимбера–Мичела	0.706	Шапиро–Уилка 2	0.266
7	Шапиро–Уилка 2	0.290	Шапиро–Уилка 2	0.657	<b>Дешпанде (0.1)</b>	0.241
8	<b>Дешпанде (0.3)</b>	0.194	Шапиро–Уилка 1	0.621	Шапиро–Уилка 1	0.223
9	Бартлетта–Морана	0.143	Корреляционный критерий 2	0.311	Корреляционный критерий 1	0.053
10	Шермана	0.140	Корреляционный критерий 1	0.276	Корреляционный критерий 2	0.016

## 5. Заключение

Критерий Дешпанде обладает рядом достоинств. Благодаря наличию нормальной аппроксимации статистики (2) нет необходимости в таблице процентных точек, что значительно упрощает работу с критерием.

Критерий Дешпанде обладает достаточно высокой мощностью.

К недостаткам критерия можно отнести неопределенность с выбором параметра  $b$ , который влияет на мощность.

Проведенные исследования мощности показывают предпочтительность выбора  $b = 0.1$ , так как в этом случае критерий обладает преимуществом в мощности относительно гипотез вида  $H_1$  (с немонотонной интенсивностью отказов). Можно отметить, что и к стандартному нормальному закону распределение статистики (2) при  $b = 0.1$  сходится быстрее, чем при  $b = 0.3$ .

В то же время относительно конкурирующих гипотез  $H_2$  и  $H_3$  предпочтительность в мощности надо отдать критерию Дешпанде при  $b = 0.3$ .

В целом можно говорить о том, что использование критерия Дешпанде при проверке показательности является вполне оправданным наряду с применением таких эффективных критериев, как критерии Фросини [7] и Кимбера–Мичела [8, 9].

## Литература

1. *Deshpande J.V.* A class of tests for exponentiality against increasing failure rate average alternatives // *Biometrika*. 1983. V.70, – P. 514-518.
2. *Dipankar Bandyopadhyay, Basu A.P.* A note on tests for exponentiality by Deshpande // *Biometrika*. 1989. V. 76, №2. – P. 403-405.
3. *Рогожников А.П., Лемешко Б.Ю.* Обзор критериев показательности // Материалы XI международной конференции “Актуальные проблемы электронного приборостроения” АПЭП-2012. Т.6, Новосибирск, 2012. – С.47-55.
4. *Blinov P.Yu., Lemeshko B.Yu.* Powers of Some Tests for Exponentiality// Proceedings of the International Workshop "Applied Methods of Statistical Analysis. Nonparametric methods in Cybernetics and system Analysis". 18-22 September 2017, Krasnoyarsk. – P. 173-180.
5. *Блинов П.Ю., Лемешко Б.Ю.* Сравнительный анализ некоторых критериев проверки показательности = The comparative analysis of some exponential tests // *Обработка информации и математическое моделирование : материалы Рос. науч.-техн. конф. [Новосибирск, 25–26 апр. 2017 г.]*. – Новосибирск : СибГУТИ, 2017. – С. 6–13.
6. *Кобзарь А. И.* Прикладная математическая статистика. Для инженеров и научных работников / А. И. Кобзарь. М : ФИЗМАТЛИТ, 2006. – 816 с.
7. *Frosini B. V.* On the distribution and power of a goodness-of-fit statistic with parametric and nonparametric applications, „Goodness-of-fit” / Ed. by Revesz P., Sarkadi K., Sen P. K., Amsterdam-Oxford-New York: North-Holland. Publ. Comp., 1987, – P. 133-154.
8. *Kimber A. C.* Tests for the exponential, Weibull and Gumbel distributions based on the stabilized probability plot // *Biometrika*. 1985. V. 72, №3. – P. 661-663.
9. *Michael J. R.* The stabilized probability plot // *Biometrika*. 1983. V. 70. – P. 11-17.

### **Блинов Павел Юрьевич**

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

### **Лемешко Борис Юрьевич**

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

## **The properties of the tests for exponentiality by Deshpande**

**P. Yu. Blinov, B. Yu. Lemeshko**

The properties of test Deshpande intended for testing of exponentiality have been considered. Distributions of test statistic, the power of test under different competing hypotheses with dependence from parameter  $b$  have been studied. The power of test is compared with powers of other exponential tests. We provide recommendations for using test based on results of studies.

*Keywords:* exponential distribution, testing hypothesis, power of test.