



УДК 519.233.3

DOI: <http://dx.doi.org/10.21686/2500-3925-2024-6-57-67>

А.А. Типикин, А.А. Прусаков, Н.А. Тимошенко

Военный учебно-научный центр Военно-морского флота
«Военно-морская академия имени Адмирала Флота Советского Союза
Н.Г. Кузнецова, Санкт-Петербург, Россия

Программная реализация критерия Эппса-Палли в среде моделирования Matlab

Цель работы. Системы моделирования и платформы программирования дают широкие возможности по использованию статистических инструментов в научно-исследовательской деятельности. Так как нормальное распределение является одним из наиболее часто встречающихся законов распределения, то критерий проверки выборки на нормальность имеет высокую востребованность среди инструментов статистического оценивания, среди которых критерий Эппса-Палли имеет статус одного из наиболее мощных. тестов для проверки отклонения распределения от нормального. Есть ряд реализаций данного теста на языках R и Python. Однако в одной из наиболее популярных систем моделирования Matlab данный тест не реализован. Таким образом, цель данного исследования разработать программную реализацию критерия Эппса-Палли в среде Matlab и проверить корректность проводимых расчетов.

Материалы и методы. Вычисление статистики Эппса-Палли реализовано двумя методами – классическим, с помощью циклов, и матрично-векторным – с помощью операций линейной алгебры. Классический метод требует вычисления промежуточных величин, необходимых для получения статистики критерия, с помощью двух независимых циклов, причем второй цикл является двойным, в котором один цикл вложен в другой. Матрично-векторный метод требует меньше строк кода за счет выполнения расчетов с помощью операций линейной алгебры над матрицами и векторами. Критические значения статистики для объема выборки от 8 до 1000 элементов получены с помощью

двухмерной линейной интерполяции табличных значений. Для выборки свыше 1000 элементов использована аппроксимация бета-функцией III рода.

Результаты. Оценка вычислительной эффективности способов показала, что циклический подход примерно в три раза превосходит матрично-векторный по показателю затрачиваемого времени, что предположительно связано с обработкой незначительных элементов в треугольных матрицах при выполнении покомпонентных операций. Корректность программной реализации теста Эппса-Палли проверена на нескольких примерах, которые подтвердили соответствие рассчитываемых значений статистики критерия, а также критических значений статистики, известным данным. Проведена статистическая оценка критерия по эмпирическим значениям ошибки первого рода. Получено соответствие значений ошибок задаваемым уровням значимости. Проведены сравнительные оценки критерия Эппса-Палли с критериями Андерса-Дарлинга и Шапиро-Уилка по показателю эмпирической мощности критерия. Результаты оценок табулированы.

Программная реализация критерия Эппса-Палли опубликована на Интернет-ресурсе MATLAB Central и доступна для свободного использования.

Ключевые слова: критерий Эппса-Палли, нормальное распределение, моделирование, программная реализация, вычислительная эффективность, ошибка первого рода, мощность критерия.

Alexey A. Tipikin, Alexander A. Prusakov, Nikolai A. Timoshenko

Military Training and Research Center of the Navy
"Naval Academy named after Admiral of the Fleet of the Soviet Union N.G. Kuznetsov", St. Petersburg, Russia

Software Implementation of the Epps-Pulley Criterion in Matlab Modeling Environment

Purpose. Modeling systems and programming platforms provide broad opportunities for the use of statistical tools in research activities. Since the normal distribution is one of the most common distribution laws, the criterion for checking the sample for normality is in high demand among statistical assessment tools, among which the Epps-Pulley test has the status as one of the most powerful tests to check the deviation of the distribution from the normal one. There are a number of implementations of this test in the R and Python languages. However, this test is not implemented in one of the most popular Matlab modeling software. Thus, the purpose of this study is to develop a software implementation of the Epps-Pulley criterion in the Matlab environment and verify the correctness of the performed calculations.

Materials and methods. We implemented the calculation of Epps-Pulley statistics by two methods – classical, using cycles, and matrix-vector, using linear algebra operations. The classical method requires calculating the intermediate values necessary to obtain the criterion statistics using two independent cycles, the second cycle being a double one, in which one cycle is nested within the other. The matrix-vector method requires fewer lines of code by performing calculations using linear algebra operations on matrices and vectors. We obtained critical statistical values for the sample size ranging from 8 to 1000 elements with two-dimensional linear interpolation

of tabular values. We used an approximation by a beta function of the third kind for a sample of over 1000 elements.

Results. An assessment of the computational efficiency of the methods showed that the cyclic approach is about three times higher than the matrix-vector approach in terms of consumed time, which is presumably associated with the processing of insignificant elements in triangular matrices when performing component-by-component operations. The correctness of the software implementation of the Epps-Pulley criterion was tested on several examples, which confirmed the compliance of the calculated values of the criterion statistics, as well as the critical values of statistics, with known data. We carried out a criterion statistical evaluation based on the empirical values of the error of the first kind. We obtained the error values correspondence to the specified significance levels. We performed comparative estimates of the Epps-Pulley test with the Anders-Darling and Shapiro-Wilk tests in terms of the criterion empirical power. Evaluation results are tabulated. The software implementation of the Epps-Pulley test is published on the MATLAB Central Internet resource and is available for free use.

Keywords: Epps-Pulley criterion, normal distribution, simulation, software implementation, calculation efficiency, error of the first kind, criterion power.

Введение

Проверка распределения погрешностей (остатков) моделирования на нормальность является неотъемлемой частью проверки практически любой разработанной модели на адекватность [1, 2]. В настоящее время существует множество критериев оценки нормальности распределения, среди которых как широко известные критерии Пирсона и Колмогорова [3, 4], так и более современные Андерса-Дарлинга, Шапиро-Уилка и Эппса-Палли [5–7] и др.

В работах [8, 9] установлено, что критерий Эппса-Палли является одним из наиболее мощных. В частности, по этому показателю он превосходит критерии Шапиро-Уилка, Д’Агостино, Дэвида-Хартли-Пирсона, однако имеет недостаток – при малых объемах выборки неспособен отличать от нормального закона, распределения с более плоскими плотностями вероятностей (с коэффициентом эксцесса $E_s \leq 2$); но этот недостаток также свойственен и критерию Шапиро-Уилка [10]. Ряд достоинств критерия Эппса-Палли описан в работе [7].

В то же самое время многие научные и инженерные системы компьютерной математики, используемые в интересах моделирования, имеют в своем составе пакеты статистического анализа, которые позволяют интегрировать процесс статистической проверки моделей. В достаточно популярной системе моделирования Matlab при наличии пакета расширения Statistics and Machine Learning для проверки нормальности распределения предлагается использовать критерии Пирсона, Колмогорова-Смирнова и Лиллиефорса [11]. Критерий Шапиро-Уилка реализован пользовательским сообществом [12], однако критерий Эппса-Палли в данной системе отсутствует, при том,

что он реализован на языках Python и R [13–15].

Таким образом, целью работы является расширение возможностей по использованию современных критериев проверки на нормальность в системах компьютерного моделирования путем реализации критерия Эппса-Палли в системе моделирования Matlab.

Методы расчета статистики критерия Эппса-Палли

Порядок расчета статистики критерия Эппса-Палли приведен в [16, 17] и общем выражается формулой:

$$T_{EP} = 1 + \frac{N}{\sqrt{3}} + B - A, \quad (1)$$

где

$$A = \sqrt{2} \sum_{j=1}^N \exp\left(-\frac{(x_j - \bar{x})^2}{4m_2}\right); \quad (2)$$

$$B = \frac{2}{N} \sum_{k=2}^N \sum_{j=1}^{N-1} \exp\left(-\frac{(x_j - \bar{x})^2}{2m_2}\right), \quad (3)$$

В формулах (1)–(3) введены обозначения N – объем выборки, \bar{x} – среднее выборочное, m_2 – центральный момент 2-го порядка.

Как видно из формул (2) и (3) величины A и B вычисляются с помощью многократного суммирования, что обычно традиционно реализуется в виде цикла [13, 18]. Алгоритм расчета величин A и B с помощью циклов показан на рис. 1.

Если учитывать, что система моделирования Matlab изначально оптимизирована для выполнения матричных операций, то логично предположить, что избавление от циклов и переход к операциям линейной алгебры повысит вычислительную эффективность алгоритма.

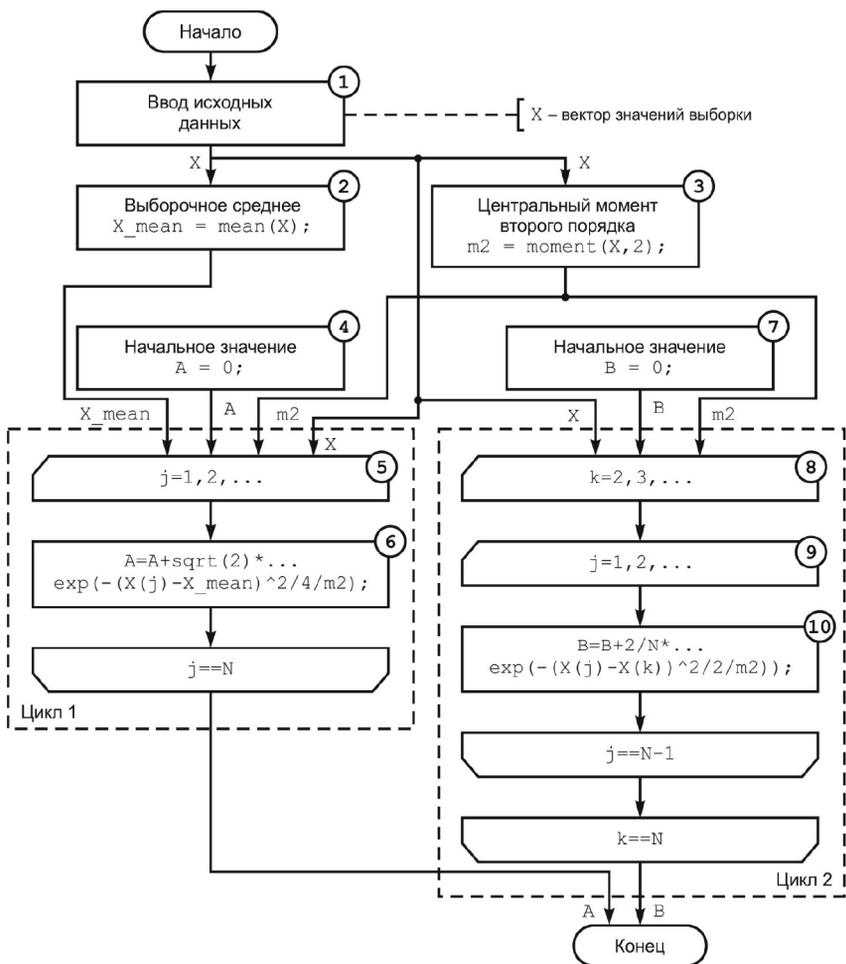


Рис. 1. Алгоритм расчета величин A и B с помощью циклов
Fig. 1. The algorithm for calculating A and B values using cycles

Тогда величины A и B могут быть рассчитаны по формулам:

$$A = \sqrt{2} \sum \left\langle \exp \left(-\frac{(X - \bar{X})^{(2)}}{4m_2} \right) \right\rangle;$$

$$B = \frac{2}{N} \sum \left\langle L \left\langle \exp \left(-\frac{(X_j - X_k)^{(2)}}{2m_2} \right) \right\rangle \right\rangle,$$

где

$$\mathbf{Xk} = (\mathbf{U}_1 \langle \times \rangle \mathbf{X}_1)^T;$$

$$\mathbf{Xj} = \mathbf{L}_1 \langle \times \rangle \mathbf{X}_2;$$

$$\mathbf{X} = \{x_1, x_2, \dots, x_N\};$$

$$\mathbf{X}_1 = \{x_2, x_3, \dots, x_N\};$$

$$\mathbf{X}_2 = \{x_1, x_2, \dots, x_{N-1}\},$$

здесь обозначены: \mathbf{U}_1 – верхняя треугольная матрица размера $(N - 1) \times (N - 1)$, все ненулевые элементы которой равны единице, \mathbf{L}_1 – нижняя треугольная матрица размера $(N - 1) \times (N - 1)$, все ненулевые элементы которой равны единице, $\mathbf{L}\langle \dots \rangle$ – оператор преобразования матрицы в нижнюю треугольную матрицу (приравниваются нулю все элементы выше главной диагонали), $\Sigma\langle \dots \rangle$ – оператор суммирования всех элементов матрицы (вектора), $\langle \times \rangle$ – оператор покомпонентного умножения, $X^{(2)}$ – покомпонентное возведение в квадрат.

Следовательно, алгоритм вычисления величин A и B вместо циклов (см. рисунок 1) может быть представлен операциями над матрицами и векторами. Алгоритм вычисления величин A и B с помощью операций линейной алгебры показан на рис. 2, причем операции в блоках 4, 5 и 6 могут выполняться параллельно.

Сравнение вычислительной эффективности способов расчета статистики

Для сравнения вычислительной эффективности по критерию наименьшего затрачиваемого времени проведены расчеты вычислений статистики критерия Эппса-Палли с различными объемами выборки. Для каждой выборки выполнено по 500 расчетов. На рисунках 3 и 4 изображены

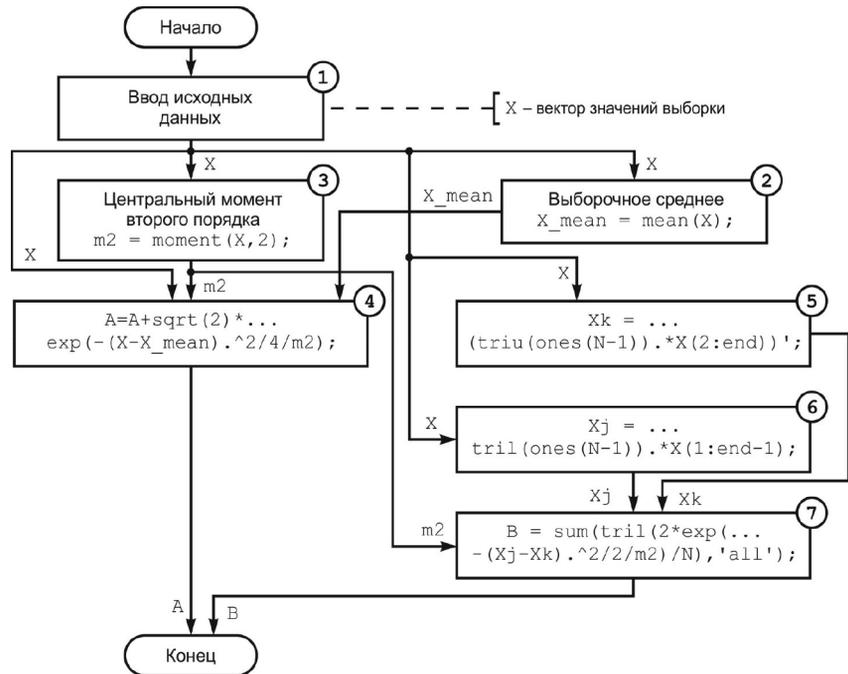


Рис. 2. Алгоритм вычисления величин A и B с помощью операций линейной алгебры

Fig. 2. The algorithm for calculating A and B values using linear algebra operations

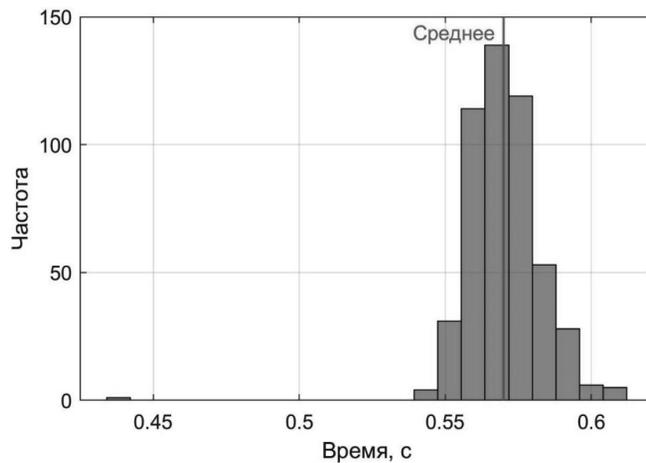


Рис. 3. Время, затрачиваемое на вычисление статистики критерия Эппса-Палли с помощью матрично-векторных операций при объеме выборки 5000 наблюдений

Fig. 3. The time consumed for the Epps-Pulley criterion statistics calculation using matrix-vector operations with a sample size of 5000 observations

гистограммы распределения времени, затрачиваемого на вычисление статистики критерия Эппса-Палли с помощью матрично-векторных операций и с помощью циклических расчетов при объеме выборки 5000 наблюдений.

Средние значения затрачиваемого времени сведены в табл. 1.

Как видно из таблицы, по показателю времени, затрачива-

емого на проведение расчетов, вариант 2 имеет приблизительно трехкратное преимущество над вариантом 1. Предположительно это объясняется дополнительным временем, связанным с выполнением операций с незначительными (нулевыми) элементами в треугольных матрицах. Возможно, это время можно уменьшить, если оптимизировать вычислительные алгоритмы, но для этого требуется

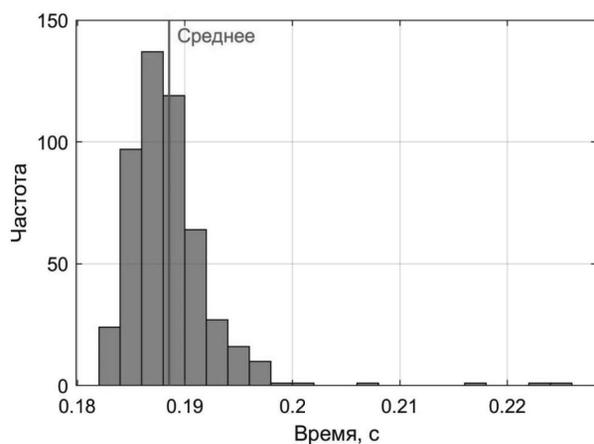


Рис. 4. Время, затрачиваемое на вычисление статистики критерия Эппса-Палли с помощью циклических расчетов при объеме выборки 5000 наблюдений

Fig. 4. The time consumed for the Epps-Pulley criterion statistics calculation using cyclic calculations with a sample size of 5000 observations

Таблица 1 (Table 1)

Средние значения времени, затрачиваемого на вычисления
Average values of the time consumed for calculations

Объем выборки	100	500	1000	5000	10000	15000
	Время, с					
Матрично-векторный способ (вариант 1)	$2,14 \cdot 10^{-4}$	$5,06 \cdot 10^{-3}$	$2,24 \cdot 10^{-2}$	0,570	2,26	4,97
Циклический способ (вариант 2)	$1,85 \cdot 10^{-4}$	$2,00 \cdot 10^{-3}$	$7,66 \cdot 10^{-2}$	0,189	0,75	1,68

низкоуровневый доступ к памяти электронно-вычислительной машины, который в Matlab не обеспечивается.

Расчет критических значений статистики

Критические значения статистики Эппса-Палли для объемов выборки от 8 до 200 и уровней значимости от 0,01 до 0,1 приведены в [16]. В работе [8] представлены данные для объемов выборки от 8 до 1000 и уровней значимости от 0,01 до 0,15. Указанные сведения могут быть использованы в табличном виде. Для получения промежуточных критических значений статистики выполняется двухмерная линейная интерполяция.

При объеме выборки более 1000 наблюдений целесообразно использовать приближенные распределения статистики критерия с помощью бета-распределения III рода с предельными параметрами [8].

Кумулятивная функция бета-распределения записывается как [4]:

$$F(x|a,b) = \frac{B(x|a,b)}{B(a,b)} = \frac{1}{B(a,b)} \int_0^x t^{a-1}(1-t)^{b-1} dt; \quad x \in [0,1], a > 0, b > 0, \quad (4)$$

где бета-функция имеет вид:

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = \int_0^1 t^{a-1}(1-t)^{b-1} dt,$$

здесь $\Gamma(z)$ – гамма-функция, выражаемая формулой:

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt. \quad (5)$$

Если в формуле (4) вместо x подставить производящую функцию

$$g(x|\delta, \alpha, \beta) = \frac{\delta^{\frac{x-\alpha}{\beta}}}{1 + (\delta-1) \frac{x-\alpha}{\beta}},$$

то получим бета-распределение III рода, аппроксимирую-

щее распределение статистики Эппса-Палли:

$$F(x|\theta_0, \theta_1, \theta_2, \theta_3, \theta_4) = \frac{B(g(x|\theta_2, \theta_3, \theta_4)|\theta_0, \theta_1)}{B(\theta_0, \theta_1)}, \quad (6)$$

с предельными значениями параметров [8]:

$$\theta_0 = 1,7669; \theta_1 = 2,1668; \theta_2 = 6,7594; \theta_3 = 0,91; \theta_4 = 0,0016.$$

Графики кумулятивной функции ($F(x)$) и плотности вероятности ($f(x)$) бета-распределения III рода, полученные по формуле (6) с предельными параметрами, показаны на рис. 5.

Оценка корректности реализации критерия

Код функции `eptest`, выполняющей проверку нормальности по критерию Эппса-Палли доступен на ресурсе Matlab-Central [19]. Данная функция возвращает результат проверки, расчетную величину статистики и критическое значение статистики Эппса-Палли. Если расчетное значение статистики не превышает критическое, то результат проверки равен 0 (нулевая гипотеза не отклоняется на заданном уровне значимости), в ином случае – 1 (нулевая гипотеза отклоняется на заданном уровне значимости).

Корректность тестирования гипотез проверена на нескольких примерах. Один из них – применение критерия Эппса-Палли для проверки нормальности распределения результатов измерения прочности вискозной нити [16]. Нормальный вероятностный график измеренных значений показателя прочности вискозной нити показан на рис. 6.

Как видно из рисунка, данные измерений сильно отклоняются от нормального распределения в хвостах. Это дает основание предполагать о несоответствии закона распределения выборки нормальному закону. Тестирование по кри-

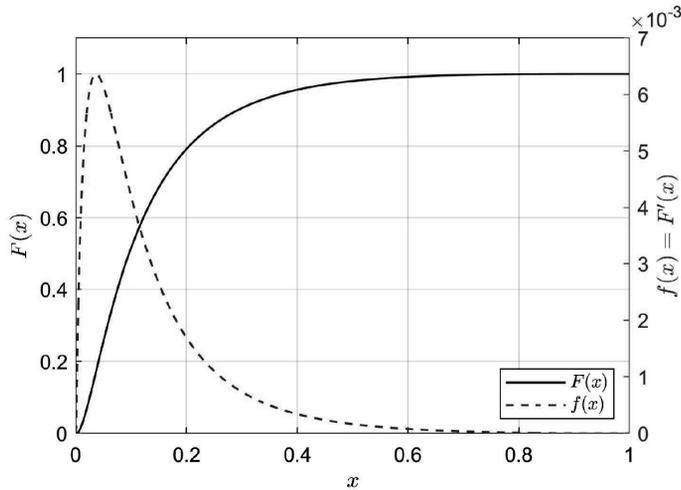


Рис. 5. Графики кумулятивной функции (F(x)) и плотности вероятности (f(x)) бета-распределения III рода с предельными параметрами

Fig. 5. Graphs of the cumulative function (F(x)) and probability density (f(x)) of the beta distribution of the third kind with limiting parameters

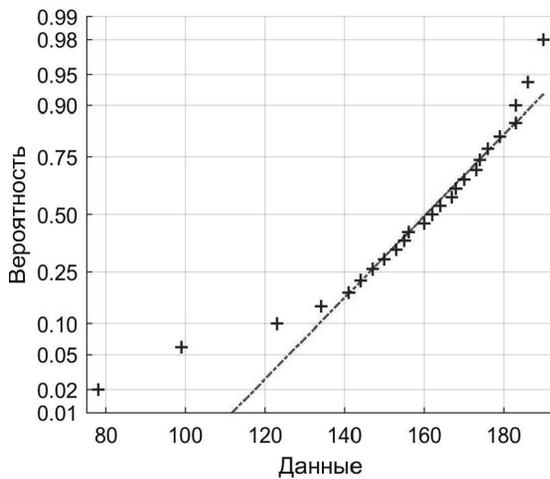


Рис. 6. Нормальный вероятностный график значений показателя прочности вискозной нити

Fig. 6. Normal probabilistic graph of the values of the viscose filament strength index

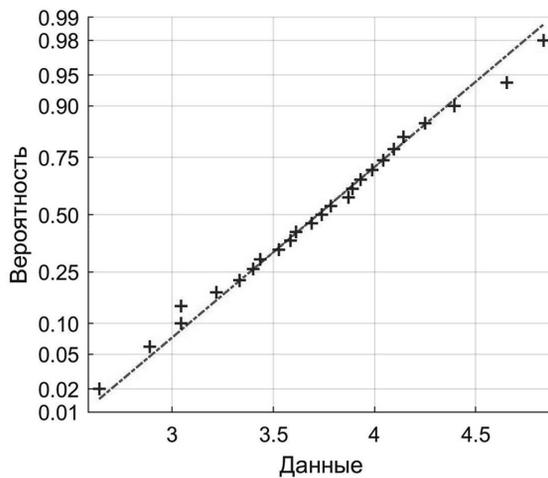


Рис. 7. Нормальный вероятностный график преобразованных значений показателя прочности вискозной нити

Fig. 7. Normal probabilistic graph of the transformed values of the viscose filament strength index

терию Эппса-Палли отвергает нулевую гипотезу на уровне значимости 0,01. При этом рассчитанное значение статистики 0,612 совпадает с данными примера. Однако критическое значение 0,57 отличается, что связано с различиями в таблицах критических значений в [8] и [16].

Если выполнить логарифмическое преобразование значений выборки по формуле $\ln(204 - x)$, то преобразованные значения показателя прочности вискозной нити гораздо лучше совпадают с прямой нормального распределения, как это показано на рис. 7.

Теперь тестирование по критерию Эппса-Палли не отвергает гипотезу о нормальном распределении данных. Расчетное значение статистики составляет 0,006, что также совпадает с примером.

Другой пример представляет собой данные измерений скорости света, выполненные С.Ньюкомом [20]. Гистограмма результатов измерений показана на рис. 8.

Как видно из рис. 8 в массиве данных присутствуют выбросы, искажающие результаты тестирования. В этом случае гипотеза о нормальности распределения отвергается на уровне значимости 0,01. Рассчитанное значение статистики 4,12 превышает критическое значение, равное 0,58.

Если из выборки убрать все величины, выходящие за трехкратное абсолютное отклонение от медианы, рассчитываемое по формуле:

$$MAD = c \cdot \text{med}(|x_i - \text{med}(\mathbf{X})|),$$

где

$$c = -\frac{1}{\sqrt{2}} \text{erfcinv} \frac{3}{2};$$

erfcinv — обратная дополнительная функция ошибок, то повторное тестирование не отклоняет нулевую гипотезу на уровне значимости 0,01, при этом рассчитанное значение статистики составляет 0,11.

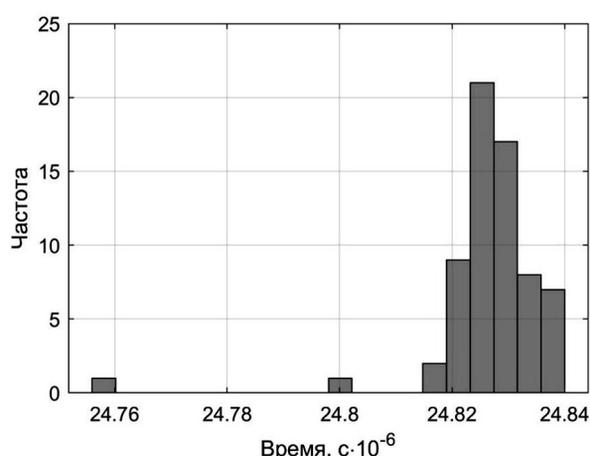


Рис. 8. Гистограмма результатов измерений скорости света, выполненных С. Ньюкомом

Fig. 8. Histogram of light speed measurements results, performed by S. Newcomb

Таблица 2 Table 2

Эмпирические вероятности ошибки первого рода для критериев Андерса-Дарлинга, Шапиро-Уилка и Эппса-Палли

Empirical probabilities of error of the first kind for the Anders-Darling, Shapiro-Wilk and Epps-Pulley tests

Объем выборки	Уровень значимости				
	0,15	0,1	0,05	0,025	0,01
10	0,1521	0,0999	0,053	0,0257	0,0114
	0,1783	0,1185	0,0632	0,0316	0,0133
	0,1496	0,0977	0,0513	0,0268	0,0118
20	0,1544	0,0981	0,0525	0,0256	0,0117
	0,182	0,1227	0,0613	0,0311	0,0125
	0,1469	0,1004	0,0531	0,0252	0,0113
30	0,1566	0,1036	0,0478	0,024	0,0102
	0,186	0,1235	0,06	0,0285	0,0106
	0,1514	0,0979	0,0465	0,0232	0,0096
40	0,147	0,0996	0,0482	0,0241	0,0093
	0,1785	0,1237	0,0621	0,0306	0,0112
	0,1459	0,101	0,0501	0,0255	0,01
50	0,1451	0,1074	0,0446	0,0261	0,0106
	0,1762	0,1301	0,0563	0,0321	0,0118
	0,1437	0,1045	0,0452	0,028	0,0122
100	0,1442	0,1008	0,0481	0,0257	0,0089
	0,1738	0,12	0,0589	0,0302	0,0127
	0,1484	0,1032	0,0481	0,0267	0,0101
150	0,1548	0,094	0,0487	0,0256	0,0085
	0,1751	0,1137	0,0627	0,0295	0,011
	0,1529	0,0921	0,0488	0,0242	0,0089
300	0,1501	0,099	0,053	0,0256	0,0106
	0,1718	0,1163	0,0628	0,0305	0,0121
	0,1503	0,1016	0,0544	0,0243	0,0101
MAE	0,0040	0,0027	0,0026	0,0008	0,0010
MARE	0,0277	0,0211	0,0109	0,0055	0,0019
	0,0025	0,0029	0,0025	0,0013	0,0009
	0,03	0,03	0,05	0,03	0,10
MARE	0,18	0,21	0,22	0,22	0,19
	0,02	0,03	0,05	0,05	0,09

Кроме совпадений на результатах отдельных примеров интерес представляет статистическая оценка результатов тестирования. Для этого взяты по 10000 выборок (повторений) с нормальным распределением $N(0, 1)$ объемами {10; 20; 30; 40; 50; 100; 150; 300} наблюдений. Для каждой выборки проведено тестирование с помощью критериев Андерса-Дарлинга [11], Шапиро-Уилка [12] и Эппса-Палли [19]. Эмпирическая вероятность ошибки первого рода должна приблизительно соответствовать заданному уровню значимости. В табл. 2 приведены полученные вероятности ошибки первого рода для указанных критериев при уровнях значимости {0,15; 0,1; 0,05; 0,025; 0,01}.

В табл. 2 в каждой ячейке для заданных уровня значимости и объема выборки указаны три эмпирических значения вероятности ошибки первого рода, верхняя из которых соответствует результатам тестирования по критерию Андерса-Дарлинга, средняя – Шапиро-Уилка и нижняя – Эппса-Палли. Внизу таблицы приведены значения средних абсолютных погрешностей, рассчитанных по формуле:

$$MAE = \frac{1}{N} \sum_{i=1}^N |x_j - \alpha|,$$

где N – количество учитываемых выборок разного объема;
 α – уровень значимости, и средних относительных погрешностей:

$$MARE = \frac{MAE}{\alpha}.$$

Из табл. 2 видно, что критерии Андерса-Дарлинга и Эппса-Палли имеют приблизительно равные погрешности при определении вероятности ошибки первого рода, в то время как аналогичные погрешности для критерия Шапиро-Уилка существенно выше, что может свидетельствовать о неточностях при

программной реализации данного критерия.

Эмпирическая мощность критерия относительно некоторых конкурирующих гипотез

Одной из наиболее важных для критерия Эппса-Палли конкурирующих гипотез является обобщенное нормальное распределение, функция плотности распределения вероятностей которого имеет вид [8]:

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

где гамма-функция $\Gamma(z)$ выражается формулой (5).

Также в качестве конкурирующих гипотез имеет смысл рассматривать логистическое распределение и распределение Лапласа, определяемые соответствующими функциями плотностей распределения вероятностей [8, 21]:

$$f(x) = \frac{1}{2\theta_1} \exp\left(-\frac{|x-\theta_0|}{\theta_1}\right);$$

$$f(x) = \frac{\frac{\pi}{\theta_1\sqrt{3}} \exp\left(-\frac{\pi(x-\theta_0)}{\theta_1\sqrt{3}}\right)}{\left(1 + \exp\left(-\frac{\pi(x-\theta_0)}{\theta_1\sqrt{3}}\right)\right)^2}.$$

Для представления о характере данных функций их графики показаны на рис. 9.

На рис. 9 изображено стандартное нормальное распределение $\mathcal{N}(0, 1)$, другие распределения имеют параметры:

обобщенное нормальное:

$\theta_0 = 0, \theta_1 = 2, \theta_2 = 4;$

Лапласа: $\theta_0 = 0, \theta_1 = 1;$

логистическое: $\theta_0 = 0, \theta_1 = 0,7.$

Методом отсечки по уровню [22] сгенерированы по 10000 выборок объемами {10; 20; 30; 40; 50; 100; 150; 300} наблюдений. Для уровней значимости {0,15; 0,1; 0,05; 0,025; 0,01} проведено тестирование каждой выборки с помощью критериев Андерса-Дарлинга [11], Шапиро-Уилка [12] и Эппса-Палли [19]. Доля прове-

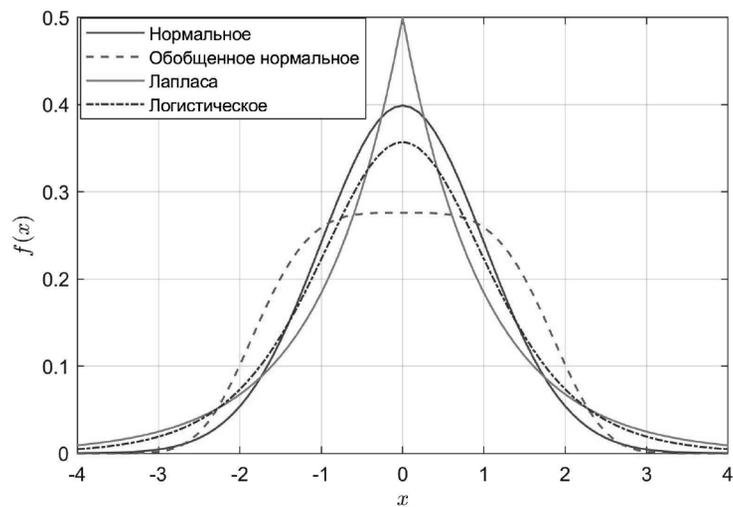


Рис. 9. Графики функций плотности распределения вероятностей
Fig. 9. Graphs of probability density functions

рок, отклонившая ложную гипотезу, должна приблизительно соответствовать мощности критерия. Полученные результаты для критерия Эппса-Пал-

ли в целом соответствуют сведениям, приведенным в [8]. Наибольшее расхождение (до 5%) получено при оценке мощности с конкурирующей гипо-

Таблица 3 (Table 3)

Эмпирические значения мощности критериев Андерса-Дарлинга, Шапиро-Уилка и Эппса-Палли при выборе обобщенного нормального распределения в качестве конкурирующей гипотезы
Empirical power values of the Anders-Darling, Shapiro-Wilk, and Epps-Pulley tests with the generalized normal distribution as a competing hypothesis

Объем выборки	Уровень значимости				
	0,15	0,1	0,05	0,025	0,01
10	0,1818	0,1184	0,0583	0,0271	0,009
	0,1966	0,1273	0,0612	0,0262	0,0076
	0,1682	0,1024	0,0446	0,0174	0,0056
20	0,2152	0,1507	0,0788	0,041	0,0146
	0,2356	0,1593	0,0736	0,031	0,0096
	0,2114	0,139	0,0601	0,0263	0,0082
30	0,2707	0,1925	0,1023	0,0545	0,0247
	0,2949	0,2041	0,1014	0,0434	0,0128
	0,2808	0,1837	0,0862	0,0356	0,0125
40	0,3264	0,2347	0,1388	0,0809	0,0341
	0,3594	0,2511	0,1346	0,0685	0,0225
	0,3467	0,2422	0,1212	0,06	0,0174
50	0,3722	0,2853	0,1886	0,1058	0,0487
	0,4126	0,3065	0,1813	0,0933	0,0332
	0,4068	0,2971	0,1713	0,0824	0,0269
100	0,6131	0,512	0,3802	0,2551	0,1474
	0,6712	0,5663	0,4028	0,2665	0,1362
	0,6654	0,5539	0,387	0,238	0,1147
150	0,7805	0,7067	0,5755	0,444	0,2957
	0,84	0,7632	0,6329	0,4833	0,3093
	0,8349	0,7541	0,6025	0,4429	0,2675
300	0,9735	0,9549	0,9074	0,8441	0,7368
	0,9873	0,9769	0,9485	0,9028	0,8092
	0,9869	0,9714	0,9312	0,8666	0,7445

Таблица 4 (Table 4)

Эмпирические значения мощности критериев Андерса-Дарлинга, Шапиро-Уилка и Эппса-Палли при выборе распределения Лапласа в качестве конкурирующей гипотезы

Empirical power values of the Anders-Darling, Shapiro-Wilk, and Epps-Pulley tests with choosing the Laplace distribution as a competing hypothesis

Объем выборки	Уровень значимости				
	0,15	0,1	0,05	0,025	0,01
10	0,2757	0,2092	0,1409	0,0918	0,055
	0,3004	0,2338	0,1565	0,1025	0,057
	0,2515	0,1952	0,1271	0,0863	0,0485
20	0,3947	0,3245	0,2436	0,181	0,1236
	0,4248	0,3505	0,2665	0,1974	0,1302
	0,3498	0,2881	0,2165	0,161	0,1071
30	0,4878	0,4229	0,3255	0,2622	0,187
	0,5191	0,4562	0,3518	0,2833	0,2005
	0,4391	0,3825	0,2881	0,2325	0,1603
40	0,5641	0,4987	0,4041	0,3342	0,258
	0,5954	0,5302	0,4288	0,3606	0,2734
	0,5225	0,4529	0,3657	0,2958	0,2236
50	0,6438	0,5715	0,4821	0,3933	0,3253
	0,6651	0,5979	0,5	0,4103	0,3402
	0,598	0,5316	0,4328	0,3494	0,2843
100	0,8436	0,8018	0,7358	0,6649	0,5808
	0,8556	0,8152	0,7454	0,678	0,5899
	0,8218	0,7749	0,6952	0,6249	0,5356
150	0,9372	0,9128	0,8752	0,8225	0,756
	0,9374	0,9174	0,8825	0,8246	0,7607
	0,9237	0,8961	0,8544	0,7881	0,7159
300	0,9964	0,9931	0,9901	0,9825	0,9683
	0,9959	0,9914	0,9895	0,9815	0,9669
	0,995	0,9912	0,9859	0,9773	0,9577

Таблица 5 (Table 5)

Эмпирические значения мощности критериев Андерса-Дарлинга, Шапиро-Уилка и Эппса-Палли при выборе логистического распределения в качестве конкурирующей гипотезы

Empirical power values of the Anders-Darling, Shapiro-Wilk, and Epps-Pulley tests with choosing the logistic distribution as a competing hypothesis

Объем выборки	Уровень значимости				
	0,15	0,1	0,05	0,025	0,01
10	0,192	0,1411	0,0786	0,0503	0,0233
	0,2292	0,174	0,0987	0,0631	0,0275
	0,1925	0,1486	0,084	0,0532	0,0245
20	0,2405	0,1713	0,1029	0,0634	0,035
	0,2929	0,2212	0,1462	0,0982	0,0535
	0,2361	0,1707	0,1072	0,0683	0,0364
30	0,2611	0,204	0,1247	0,083	0,0438
	0,3377	0,2707	0,1882	0,1294	0,0785
	0,2668	0,2049	0,1277	0,085	0,0481
40	0,2883	0,2211	0,1433	0,096	0,0475
	0,3826	0,3066	0,2206	0,1568	0,0964
	0,2888	0,2239	0,1487	0,1005	0,0507
50	0,3125	0,2472	0,1611	0,1092	0,0557
	0,4057	0,348	0,2475	0,1839	0,1183
	0,3148	0,2511	0,1671	0,1151	0,059
100	0,4154	0,3401	0,2417	0,1653	0,0994
	0,5473	0,4738	0,3789	0,2916	0,2093
	0,4282	0,3496	0,2489	0,1745	0,1063
150	0,5136	0,4226	0,3218	0,2224	0,1536
	0,6509	0,5795	0,4794	0,3864	0,2958
	0,5375	0,4414	0,3404	0,235	0,1591
300	0,7175	0,6508	0,5329	0,4289	0,3132
	0,8344	0,7885	0,7026	0,6263	0,5186
	0,7559	0,6815	0,5673	0,4574	0,334

тезой в виде распределения Лапласа. Однако, данная разница объясняется, во-первых, эмпирическим характером полученных значений мощности критерия, причем точность полученных оценок в работе [8] не указана; во-вторых, отсутствием данных в источнике [8] о конкретных значениях параметров распределений конкурирующих гипотез.

Поэтому дальнейшее сравнение возможно провести только между реализованными (эмпирическими) значениями мощности различных критериев, полученных способом, указанным выше. Количественные оценки представлены в табл. 3, 4 и 5. В качестве примера на рис. 10 и 11 показана относительная разница эмпирических мощностей критериев Андерса-Дарлинга и Шапиро-Уилка относительно критерия Эппса-Палли при конкурирующей гипотезе в виде обобщенного нормального распределения, которая рассчитана по формуле:

$$RE = \frac{x - x'}{x}$$

где x – эмпирическая мощность критерия Эппса-Палли для заданного количества повторений при определенных уровне значимости и объеме выборки, а x' – эмпирическая мощность критерия Андерса-Дарлинга или Шапиро-Уилка соответственно.

Из рис. 10 видно, что критерий Эппса-Палли уступает критерию Андерса-Дарлинга по показателю эмпирической мощности при малых выборках и уровнях значимости. При больших выборках и более высоких уровнях значимости критерий Эппса-Палли незначительно выигрывает. Из рис. 11 видно, что критерий Эппса-Палли во всех случаях проигрывает критерию Шапиро-Уилка, однако следует иметь в виду недостатки, реализации данного критерия, связанные с несо-

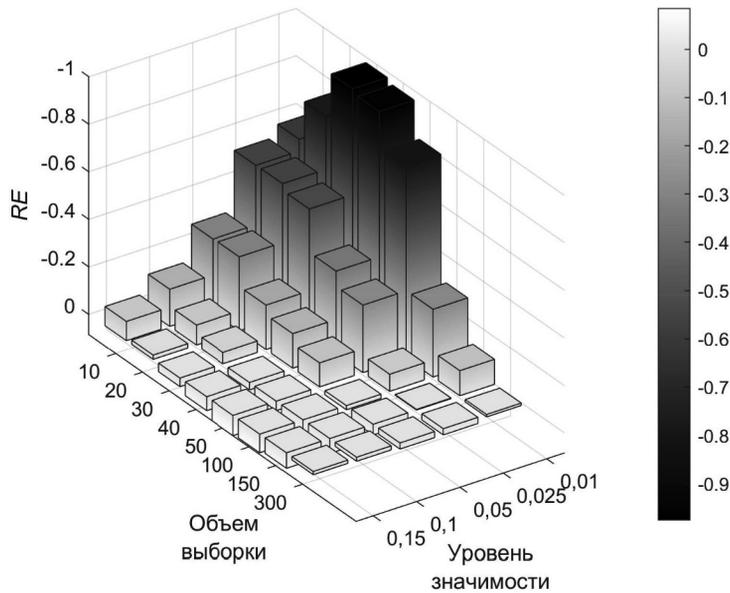


Рис. 10. Относительная разница эмпирических мощностей критериев Андерса-Дарлинга и Эппса-Палли
Fig. 10. The relative difference in the empirical power of the Anders-Darling and Epps-Pulley tests

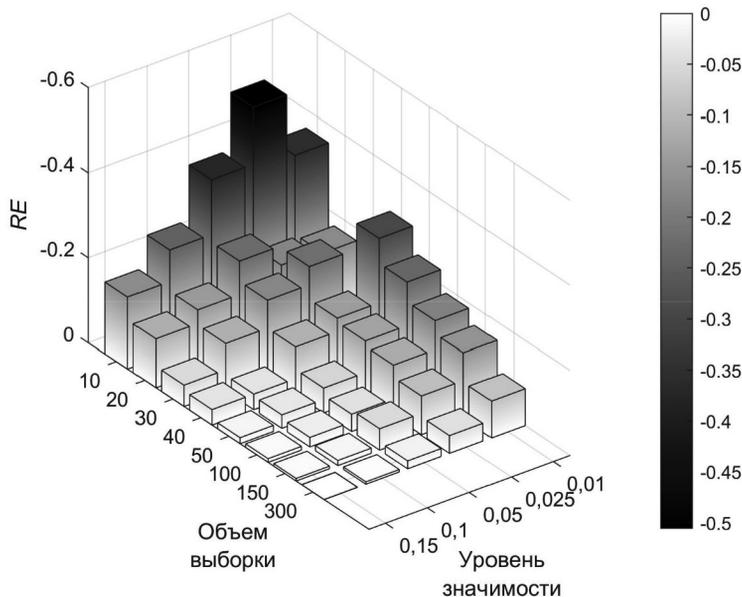


Рис. 11. Относительная разница эмпирических мощностей критериев Шапиро-Уилка и Эппса-Палли
Fig. 11. The relative difference in the empirical power of the Shapiro-Wilk and Epps-Pulley tests

ответствием эмпирической вероятности ошибки первого рода заданному уровню значимости, что было описано выше. В целом из таблиц 3, 4 и 5 можно сделать выводы, что критерии Андерса-Дарлинга и Эппса-Палли дают

соизмеримые результаты. Критерий Шапиро-Уилка по показателю эмпирической мощности превосходит критерий Эппса-Палли, однако данный результат может быть подвергнут сомнению в силу вышеуказанных причин.

Выводы

Разработанная функция `eptest` позволяет выполнять проверку гипотез о нормальности распределения выборки в среде моделирования Matlab на основе одного из наиболее мощных критериев. Исследована вычислительная эффективность двух методов расчета статистики критерия Эппса-Палли. Выявлено, что классический способ с применением циклов, примерно в три раза превосходит матрично-векторный по показателю времени, затрачиваемого на расчеты. Реализация процедуры тестирования дает корректные результаты на известных примерах, расчетные и критические значения статистики совпадают. Статистическая оценка результатов тестирования показала, что вероятность ошибки первого рода соответствует заданным уровням значимости, а погрешность оценки сравнима с реализованным критерием Андерса-Дарлинга. Сравнительная оценка критериев Андерса-Дарлинга, Шапиро-Уилка и Эппса-Палли по показателю эмпирической мощности показала примерную соизмеримость результатов для критериев Андерса-Дарлинга и Эппса-Палли. Критерий Шапиро-Уилка превзошел по мощности критерии Андерса-Дарлинга и Эппса-Палли для всех рассмотренных конкурирующих гипотез, однако в силу наличия расхождений между эмпирической ошибкой первого рода и заданными уровнями значимости для данного критерия, его превосходство может быть поставлено под сомнение. Программная реализация критерия Эппса-Палли опубликована на Интернет-ресурсе MATLAB Central и доступна для свободного использования [19].

Литература

1. Гнатюк В.И. Закон оптимального построения техноценозов. Калининград: КИЦ «Техноценоз», 2019. 940 с.
2. Murray-Smith D.J. Testing and Validation of Computer Simulation Models. Principles, Methods and Applications. New York: Springer, 2015. 258 с. DOI: 10.1007/978-3-319-15099-4.
3. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. 3-е изд. М.: Наука, 1983. 416 с.
4. Кобзарь А. И. Прикладная математическая статистика. Для инженеров и научных работников. М.: Физматлит, 2006. 816 с.
5. Волчихин В.И., Иванов А.И., Безяев А.В., Куприянов Е. Н. Нейросетевой анализ нормальности малых выборок биометрических данных с использованием хи-квадрат критерия и критериев Андерсона-Дарлинга // Инженерные технологии и системы. 2019. Т. 29. № 2. С. 205–217. DOI: 10.15507.2658-4123.029.201902.205-217.
6. Ivanov A.I., Vjatchanin S.E., Malygina E.A., Lukin V.S. Precision statistics: neuroet networking of chi-square test and Shapiro-Wilk test in the analysis of small selections of biometric data // Надежность и качество сложных систем. 2019. № 2(26). С. 27–34. DOI: 10.21685/2307-4205-2019-2-4.
7. Ebner B., Henze N. Bahadur efficiencies of the Epps–Pulley test for normality // Journal of Mathematical Sciences. 2021. № 273. С. 861–870. DOI: 10.1007/s10958-023-06547-2.
8. Лемешко Б.Ю. Критерии проверки отклонения распределения от нормального закона. Руководство по применению. Новосибирск: НГТУ, 2014. 192 с.
9. Лемешко Б.Ю. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход. Новосибирск: НГТУ, 2011. 888 с.
10. Razali N.M., Wah Y.B. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests // Journal of Statistical Modeling and Analytics. 2011. Т. 2. № 1. С. 21–33.
11. Statistics and Machine Learning Toolbox User’s Guide. [Электрон. ресурс]. Режим досту-

па: https://www.mathworks.com/help/pdf_doc/stats/stats.pdf.

12. BenSaida A. Shapiro-Wilk and Shapiro-Francia normality tests. [Электрон. ресурс]. Режим доступа: <https://www.mathworks.com/matlabcentral/fileexchange/13964-shapiro-wilk-and-shapiro-francia-normality-tests>.
13. Назаров А.А. Проверка нормальности распределения с использованием критерия Эппса-Палли средствами Python [Электрон. ресурс]. Режим доступа: <https://github.com/AANazarov/MyModulePython.git>.
14. Kelly D.E. Oceanographics Analysis with R. New York: Springer-Verlag, 2018. 280 с. DOI: 10.1007/978-1-4939-8844-0.
15. The R project for Statistical Computing [Электрон. ресурс]. Режим доступа: <https://www.r-project.org>.
16. ГОСТ Р ИСО 5479–2002. Проверка отклонения распределения вероятностей от нормального распределения. М.: Госстандарт России, 2002.
17. International Standard ISO 5479-1997. Statistical interpretation of data – Test for departure from the normal Distribution. Geneva: International Standardization Organization, 1997.
18. Altar R.R., Samanta D., Konar D., Bhattacharrya S. Software Source Code: Statistical Modeling. Berlin: De Grunter, 2021. 358 с. DOI: 10.1515/9783110703399.
19. Tipikin A.A. Epps-Pulley test for departure from normal distribution [Электрон. ресурс]. Режим доступа: <https://www.mathworks.com/matlabcentral/fileexchange/158036-eptest>.
20. Stigler S.M. Do robust estimators work with real data? // The Annals of Statistics. 1977. Т. 5. № 6. С. 1055–1098. DOI: 10.1214/aos/1176343997.
21. Бессонов А.А. Искусственный интеллект и математическая статистика в криминалистическом изучении преступлений. М.: Проспект, 2021. 816 с.
22. Karamandis M., Beutler F. Ensemble slice sampling // Statistics and Computing. 2021. Т. 31. № 61. С. 1–18. DOI: 10.1007/s11222-021-10038-2.

References

1. Gnatyuk V.I. Zakon optimal’nogo postroyeniya tekhnotsenozov = The Law of Optimal Construction of Technocenoses. Kaliningrad: KIC «Tekhnotsenoz»; 2019. 940 p. (In Russ.)
2. Murray-Smith D.J. Testing and Validation of Computer Simulation Models. Principles, Methods and Applications. New York: Springer; 2015. 258 p. DOI: 10.1007/978-3-319-15099-4.
3. Bol’shev L.N., Smirnov N.V. Tablitsy matematicheskoy statistiki. 3-ye izd. = Tables of Mathematical Statistics. 3rd ed. Moscow: Science; 1983. 416 p. (In Russ.)
4. Kobzar’ A.I. Prikladnaya matematicheskaya statistika. Dlya inzhenerov i nauchnykh rabotnikov = Applied Mathematical Statistics. For Engineers

and Researchers. Moscow: Fizmatlit; 2006. 816 p. (In Russ.)

5. Volchikhin V.I., Ivanov A.I., Bezyayev A.V., Kupriyanov Ye.N. Neural network analysis of normality of small samples of biometric data using the chi-square test and Anderson-Darling criteria. Inzhenernyye tekhnologii i sistemy = Engineering technologies and systems. 2019; 29; 2 205–217. DOI: 10.15507.2658-4123.029.201902.205-217. (In Russ.)
6. Ivanov A.I., Vjatchanin S.E., Malygina E.A., Lukin V.S. Precision statistics: neuroet networking of chi-square test and Shapiro-Wilk test in the analysis of small selections of biometric data. Nadezhnost’ i kachestvo slozhnykh system = Reliability and quality of complex systems. 2019; 2(26): 27–34. DOI: 10.21685/2307-4205-2019-2-4. (In Russ.)

7. Ebner B., Henze N. Bahadur efficiencies of the Epps–Pulley test for normality. *Journal of Mathematical Sciences*. 2021; 273: 861–870. DOI: 10.1007/s10958-023-06547-2.

8. Lemeshko B.Yu. Kriterii proverki otkloneniya raspredeleniya ot normal'nogo zakona. *Rukovodstvo po primeneniyu = Criteria for testing the deviation of a distribution from the normal law. Application Guide*. Novosibirsk: NSTU; 2014. 192 p. (In Russ.)

9. Lemeshko B.Yu. Statisticheskii analiz dannykh, modelirovaniye i issledovaniye veroyatnostnykh zakonmernostey. *Komp'yuternyy podkhod = Statistical data analysis, modeling and study of probability patterns. Computer approach*. Novosibirsk: NSTU; 2011. 888 p. (In Russ.)

10. Razali N.M., Wah Y.B. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*. 2011; 2; 1: 21–33.

11. *Statistics and Machine Learning Toolbox User's Guide* [Internet]. Available from: https://www.mathworks.com/help/pdf_doc/stats/stats.pdf.

12. BenSaida A. Shapiro-Wilk and Shapiro-Francia normality tests [Internet]. Available from: <https://www.mathworks.com/matlabcentral/fileexchange/13964-shapiro-wilk-and-shapiro-francia-normality-tests>.

13. Nazarov A.A. Proverka normal'nosti raspredeleniya s ispol'zovaniyem kriteriya Eppsa-Palli sredstvami Python = Testing the Normality of Distribution Using the Epps-Pally Test Using Python [Internet]. Available from: <https://github.com/AANazarov/MyModulePython.git>. (In Russ.)

14. Kelly D.E. *Oceanographics Analysis with R*. New York: Springer-Verlag; 2018. 280 p. DOI: 10.1007/978-1-4939-8844-0.

15. The R project for Statistical Computing [Internet]. Available from: <https://www.r-project.org>.

16. GOST R ISO 5479–2002. Proverka otkloneniya raspredeleniya veroyatnostey ot normal'nogo raspredeleniya = GOST R ISO 5479–2002. Testing the Deviation of a Probability Distribution from a Normal Distribution. Moscow: Gosstandart of Russia; 2002. (In Russ.)

17. International Standard ISO 5479-1997. Statistical interpretation of data – Test for departure from the normal Distribution. Geneva: International Standardization Organization, 1997.

18. Altar R.R., Samanta D., Konar D., Bhattacharyya S. *Software Source Code: Statistical Modeling*. Berlin: De Gruyter, 2021. 358 s. DOI: 10.1515/9783110703399.

19. Tipikin A.A. Epps-Pulley test for departure from normal distribution [Internet]. Available from: <https://www.mathworks.com/matlabcentral/fileexchange/158036-eptest>.

20. Stigler S.M. Do robust estimators work with real data? *The Annals of Statistics*. 1977; 5; 6: 1055–1098. DOI: 10.1214/aos/1176343997.

21. Bessonov A.A. *Iskusstvennyy intellekt i matematicheskaya statistika v kriminalisticheskom izuchenii prestupleniy = Artificial Intelligence and Mathematical Statistics in Forensic Study of Crimes*. Moscow: Prospekt; 2021. 816 p. (In Russ.)

22. Karamandis M., Beutler F. Ensemble slice sampling. *Statistics and Computing*. 2021; 31; 61: 1–18. DOI: 10.1007/s11222-021-10038-2.

Сведения об авторах

Алексей Алексеевич Типикин

Начальник отдела

*Военный учебно-научный центр Военно-морского флота «Военно-морская академия имени Адмирала Флота Советского Союза Н.Г. Кузнецова, Санкт-Петербург, Россия
Эл. почта: alextip@mail.ru*

Александр Александрович Прусаков

Старший научный сотрудник

*Военный учебно-научный центр Военно-морского флота «Военно-морская академия имени Адмирала Флота Советского Союза Н.Г. Кузнецова, Санкт-Петербург, Россия
Эл. почта: prusakovaa@mail.ru*

Николай Александрович Тимошенко

Младший научный сотрудник

*Военный учебно-научный центр Военно-морского флота «Военно-морская академия имени Адмирала Флота Советского Союза Н.Г. Кузнецова, Санкт-Петербург, Россия
Эл. почта: timoshenkonik@mail.ru*

Information about the authors

Alexey A. Tipikin

Head of Department

*Military Training and Research Center of the Navy “Naval Academy named after Admiral of the Fleet of the Soviet Union N.G. Kuznetsov”,
St. Petersburg, Russia
E-mail: alextip@mail.ru*

Alexander A. Prusakov

Senior Researcher

*Military Training and Research Center of the Navy “Naval Academy named after Admiral of the Fleet of the Soviet Union N.G. Kuznetsov”,
St. Petersburg, Russia
E-mail: prusakovaa@mail.ru*

Nikolai A. Timoshenko

Junior Researcher,

*Military Training and Research Center of the Navy “Naval Academy named after Admiral of the Fleet of the Soviet Union N.G. Kuznetsov”,
St. Petersburg, Russia
E-mail: timoshenkonik@mail.ru*