УДК 519.24

КРИТЕРИЙ СОГЛАСИЯ НА ОСНОВЕ БИНОМИАЛЬНОГО РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ

Е. Л. Кулешов, К. А. Петров, Т. С. Кириллова, Р. А. Халиуллин

Дальневосточный федеральный университет,
690950, г. Владивосток, ул. Суханова, 8
E-mail: kuleshov.el@dvfu.ru

На основе биномиального распределения вероятностей построен критерий согласия, который сводится к последовательности двусторонних критериев проверки гипотезы о значении функции распределения вероятностей при разных значениях её аргумента. Показано, что каждый элемент этой последовательности является несмещённым локально наиболее мощным критерием. Предложен алгоритм вычисления уровня значимости, свободный от распределения вероятностей. Качество критерия проверялось численным моделированием.

Ключевые слова: критерий согласия, интервальная оценка, закон распределения вероятностей, уровень значимости.

DOI: 10.15372/AUT20180114

Введение. Статистический анализ опытных данных часто сводится к проверке гипотезы о соответствии эмпирической функции распределения вероятностей (ФРВ) измеренной случайной величины и предполагаемой теоретической [1–5]. Для решения этой проблемы обычно используются критерии согласия, такие как критерий Колмогорова, Мизеса, хи-квадрат. Реализация этих критериев осложняется тем, что не существует точного алгоритма выбора вероятности ошибки первого рода (уровня значимости, т. е. вероятности отклонить верную гипотезу), а также тем, что вероятность ошибки второго рода (вероятность принять неверную гипотезу) не может быть вычислена априори. Например, выбор слишком малой вероятности ошибки первого рода приводит к большой вероятности ошибки второго рода и, следовательно, ухудшает качество процедуры принятия решения об истинности рассматриваемой гипотезы. Кроме этого, отметим, что различие (сходство) эмпирической функции распределения вероятностей и предполагаемой теоретической задаётся всего лишь одним числом. Единственное число в некоторых случаях может оказаться слишком общей (грубой) мерой различия двух функций вещественного аргумента. Данные особенности могут снижать качество процедуры принятия решения. В связи с этим проблема построения критериев согласия с более высокими характеристиками качества, чем известные, представляется актуальной и в литературе обсуждаются новые способы её решения [1–5].

В работе [6] предложен критерий согласия с использованием интервальной оценки функции распределения вероятностей, в котором задаётся более детальная мера различия (в виде нескольких чисел) эмпирической и теоретической ФРВ, чем единственное число. Результаты моделирования [7] показали, что для небольших размеров выборки и относительно большой дисперсии истинного распределения вероятностей по сравнению с альтернативным распределением предложенный критерий значительно эффективнее критерия Колмогорова. С ростом объёма выборки оба критерия становятся эквивалентными по качеству.

Цель предлагаемой работы — исследование нового критерия согласия в виде последовательности двусторонних критериев, построенных для разных значений аргумента ФРВ на основе биномиального распределения вероятностей. Показано, что каждый элемент последовательности является несмещённым локально наиболее мощным критерием. Предложена модификация критерия для проверки закона распределения вероятностей (ЗРВ) непрерывной случайной величины с использованием преобразования исследуемой случайной величины к величине с равномерным ЗРВ. Построен алгоритм вычисления уровня значимости, свободный от распределения вероятностей. Качество предлагаемого критерия согласия проверялось численным моделированием.

Свойства эмпирической функции распределения вероятностей. Пусть $F(x) = P(\xi \le x)$, $-\infty < x < \infty$, — истинная ФРВ исследуемой случайной величины ξ, где $P(\xi \le x)$ — вероятность события $\xi \le x$. Представим измерения величины ξ выборкой $x_1, \ldots, x_n$ размера n. Если $\nu(x_i \le x)$ — число элементов x_i выборки таких, что каждый элемент $x_i \le x$, то эмпирическая функция распределения вероятностей (точечная оценка функции F) $\hat{F}(x) = \nu(x_i \le x)/n$. Процедура измерения случайной величины ξ и вычисления оценки $\hat{F}$ непосредственно связана со следующей вероятностной схемой Бернулли. Выполняется последовательность из n опытов, в каждом из которых измеряется значение x_i случайной величины ξ. Всякое событие $\xi \le x$ будем рассматривать как успех, вероятность которого $p = P(\xi \le x) = F(x)$. Число успехов $\nu(x_i \le x)$ имеет биномиальное распределение вероятностей

$$P(\nu = k) = C_n^k p^k q^{n-k}, \quad k = 0, 1, \ldots, n, \tag{1}$$

где C_n^k — число сочетаний из n по k, а $q = 1 - p$ — вероятность неудачи, следовательно, $P(\hat{F} = k/n) = C_n^k p^k q^{n-k}$. Пусть $\mathbf{M}, \mathbf{D}$ — операторы математического ожидания и дисперсии. Тогда $\mathbf{M}\nu = np$, $\mathbf{D}\nu = npq$, $\mathbf{M}\hat{F}(x) = F(x)$ и $\mathbf{D}\hat{F}(x) = F(x)[1 - F(x)]/n$.

Несмещённый локально наиболее мощный критерий. На основе биномиального распределения вероятностей величины $\hat{F}(x)$ можно предложить следующую процедуру проверки статистической гипотезы H_0 о том, что $F_0(x)$ является истинной ФРВ. Если функция

$$\hat{F}(x) \in [z_1(x), z_2(x)], \quad x \in \{x\colon 1/n \le \hat{F}(x) \le (1 - 1/n)\}, \tag{2}$$

то гипотеза H_0 принимается, в противном случае H_0 отклоняется. В реализации алгоритма (2) аргумент x полагается дискретным и условие (2) проверяется для каждого x из области определения функции $\hat{F}(x)$. Границы $z_1(x)$, $z_2(x)$ найдём далее таким образом, чтобы для любого фиксированного x условие (2) представляло собой несмещённый локально наиболее мощный критерий заданного уровня α_0.

Пусть аргумент x фиксирован. Рассмотрим $\nu(x_i \le x)$ как выборку единичного размера случайной величины с биномиальным распределением вероятностей (1), зависимым от параметра $p = F(x)$. Тогда функция правдоподобия выборки

$$W(\nu; p) = C_n^k p^\nu (1 - p)^{n-\nu}. \tag{3}$$

Отсюда следует, что

$$\frac{\partial W(\nu; p)}{\partial p} = \frac{\partial \ln W}{\partial p} W = \frac{\nu - np}{p(1 - p)} W, \tag{4}$$

$$\frac{\partial^2 W(\nu; p)}{\partial p^2} = \Big(\frac{\partial \ln W}{\partial p}\Big)^2 W + \frac{\partial^2 \ln W}{\partial p^2} W = \frac{(\nu - np)^2 - (1 - 2p)(\nu - np) - np(1 - p)}{p^2(1 - p)^2} W. \tag{5}$$

Обозначим $p_0 = F_0(x)$, $q_0 = 1 - p_0$ и рассмотрим гипотезу $h_0 = \{p = p_0\}$ о том, что для фиксированного x параметр распределения (1) равен p_0, против альтернативы

$h_1 = \{p \neq p_0,\ 0 \leq p \leq 1\}$. Пусть D — допустимая область (если $\nu \in D$, то принимается гипотеза h_0) и K — критическая область (если $\nu \in K$, то принимается гипотеза h_1) несмещённого локального наиболее мощного критерия уровня α_0. Известно [8], что

$$D = \Big\{\nu\colon \frac{\partial^2 W(\nu,p)}{\partial p^2}\Big|_{p=p_0} < c_1 \frac{\partial W(\nu,p)}{\partial p}\Big|_{p=p_0} + c_0 W(\nu,p_0)\Big\}, \tag{6}$$

где c_1 и c_0 определяются условиями

$$\sum_{k\,\in\,K} W(k,p_0) = \alpha_0, \tag{7}$$

$$\sum_{k\,\in\,K} \frac{\partial W(k,p)}{\partial p}\Big|_{p=p_0} = 0. \tag{8}$$

Подставим выражения (3)–(5) в (6) и обозначим $z = \nu - np_0$. Тогда неравенство в соотношении (6) принимает вид

$$\frac{z^2 - (q_0 - p_0)z - np_0q_0}{p_0^2 q_0^2} < c_1 \frac{z}{p_0q_0} + c_0. \tag{9}$$

Отсюда получаем

$$z^2 - 2c_3 z < c_2, \tag{10}$$

$$2c_3 = q_0 - p_0 + c_1 p_0 q_0; \qquad c_2 = np_0q_0 + c_0 p_0^2 q_0^2. \tag{11}$$

Из (10) следует, что допустимая область

$$D = \{\nu\colon k_1 < \nu < k_2\}, \tag{12}$$

$$k_1 = np_0 + c_3 - \sqrt{c_2 + c_3^2}; \qquad k_2 = np_0 + c_3 + \sqrt{c_2 + c_3^2}. \tag{13}$$

Представим величины c_0, c_1 через k_1, k_2. Из (13) получим

$$c_3 = \frac{1}{2}\,(k_1 + k_2) - np_0, \tag{14}$$

затем из (11) будем иметь

$$c_1 = \frac{2c_3 + p_0 - q_0}{p_0q_0} = \frac{k_1 + k_2 - 2np_0 + p_0 - q_0}{p_0q_0}. \tag{15}$$

Аналогично из (13)

$$c_2 = \frac{1}{4}\,(k_2 - k_1)^2 - c_3^2, \tag{16}$$

затем из (11), (14) найдём

$$c_0 = \frac{c_2 - np_0q_0}{p_0^2q_0^2} = \frac{1}{p_0^2q_0^2}\Big[\frac{1}{4}(k_2 - k_1)^2 - \Big(\frac{k_1 + k_2}{2} - np_0\Big)^2 - np_0q_0\Big]. \tag{17}$$

Соотношения (15), (17) определяют величины c_1, c_0 через границы k_1, k_2 допустимой области. Далее представим условия (7), (8) через k_1, k_2, откуда установим границы k_1, k_2 допустимой области несмещённого локально наиболее мощного критерия. Из (7) следует

$$\sum_{k=k_1}^{k_2} W(k, p_0) = 1 - \alpha_0. \tag{18}$$

Подставим выражение (4) в (8) и после несложных преобразований получим

$$\sum_{k=k_1}^{k_2} kW(k; p_0) = n(1 - \alpha_0)p_0. \tag{19}$$

Выражения (18), (19) для заданного α_0 будем рассматривать как систему двух уравнений относительно неизвестных k_1, k_2. Её решение задаёт границы $k_1(x), k_2(x)$ допустимой области (12) случайной величины $\nu(x_i \leq x)$ для каждого аргумента x из области определения эмпирической ФРВ $\hat{F}(x)$, а также границы $z_1(x) = k_1(x)/n$, $z_2(x) = k_2(x)/n$, которые используются в алгоритме принятия решения (2). Отметим, что k_1, k_2 — целые, а правые части уравнений (18), (19) — вещественные числа. Поэтому очевидно, что система уравнений (18), (19) в общем случае не имеет решения и представляет собой некорректно поставленную задачу. Для нахождения устойчивого решения этой задачи далее используется метод минимизации сглаживающего функционала [9].

Преобразование исследуемой случайной величины к равномерному распределению. При истинной H_0 границы z_1, z_2, определяемые через решение системы уравнений (18), (19), зависят от функции $F_0(x)$, следовательно, уровень значимости α такого критерия необходимо вычислять заново для каждой гипотезы. Для устранения этого недостатка предлагается критерий согласия с использованием преобразования наблюдаемой случайной величины к величине с равномерным ЗРВ.

Пусть гипотеза H_0 верна, тогда $F_0(x)$ — ФРВ случайной величины ξ, а случайная величина $F_0(\xi)$ распределена равномерно на интервале [0, 1]. Поэтому если $x_1, \ldots, x_n$ — выборка случайной величины ξ, то $F_0(x_1), \ldots, F_0(x_n)$ — выборка равномерно распределённой на интервале [0, 1] величины $F_0(\xi)$. Сформируем для выборки $F_0(x_1), \ldots, F_0(x_n)$ вариационный ряд. Обозначим через y_i элементы этого ряда, где y_1 — минимальный, y_n — максимальный элементы, тогда

$$y_0 \leq y_1 \leq \ldots \leq y_n \leq y_{n+1}, \quad y_0 = 0, \;\; y_{n+1} = 1. \tag{20}$$

По выборке (20) построим эмпирическую ФРВ $\hat{F}_0(x)$ преобразованной случайной величины $F_0(\xi)$, при этом $\hat{F}_0(y_i) = i/n$, $i = 1, \ldots, n$. Оценка $\hat{F}_0(y)$ ФРВ растёт скачками, пропорциональными $1/n$, а соответствующая истинная ФРВ случайной величины $F_0(\xi)$ непрерывна и равна y, $y \in [0, 1]$. Максимальная погрешность оценки $\hat{F}_0(y)$ возникает в точках скачка. Для снижения таких ошибок построим новую оценку $\hat{F}_2(y)$ ФРВ случайной величины $F_0(\xi)$ в виде линейной аппроксимации, вычисляемой по значениям $\hat{F}_0(y_i)$, и определим

$$\hat{F}_1(y_i) = \frac{i - 1/2}{n}, \;\; i = 1, \ldots, n; \quad \hat{F}_1(y_0) = 0; \quad \hat{F}_1(y_{n+1}) = 1, \tag{21}$$

затем вычислим

$$\hat{F}_2(y) = \hat{F}_1(y_i) + (y - y_i)\frac{\hat{F}_1(y_{i+1}) - \hat{F}_1(y_i)}{y_{i+1} - y_i}, \quad y_i \le y \le y_{i+1}, \; i = 0, 1, \ldots, n. \tag{22}$$

Таким образом, соотношение (22) определяет функцию $\hat{F}_2(y)$ для $y \in [0, 1]$. Рассмотрим пример. Пусть y_{i+1} — точка скачка функции $\hat{F}_0(y)$ и в окрестности справа от y_{i+1} значение функции $\hat{F}_0(y)$ равно $\hat{F}_0(y_{i+1})$, а в окрестности слева — $\hat{F}_0(y_i)$. Тогда если истинная ФРВ в точке y_{i+1} равна $\hat{F}_0(y_i)$, то уклонение $\hat{F}_0(y)$ от истинной ФРВ в окрестности справа достигает величины скачка $\hat{F}_0(y_{i+1}) - \hat{F}_0(y_i)$. При этом уклонение функции $\hat{F}_2(y)$ составит половину скачка. Возможные большие уклонения от истинной ФРВ для оценки $\hat{F}_2$ снижаются на половину скачка.

Для преобразованной случайной величины $F_0(\xi)$ в соотношениях (18), (19) $p_0(x)$, $-\infty < x < \infty$, заменяется величиной y, $y \in [0, 1]$. При этом устойчивое решение k_1, k_2 системы уравнений (18), (19) для фиксированного y определим из условия минимума сглаживающего функционала

$$\Big[\sum_{k=k_1}^{k_2} C_n^k y^k (1-y)^{n-k} - (1-\alpha_0)\Big]^2 + \Big[\frac{1}{n}\sum_{k=k_1}^{k_2} k C_n^k y^{k-1} (1-y)^{n-k} - (1-\alpha_0)\Big]^2 +$$

$$+\frac{\alpha_0}{n^3}\,(k_2 - k_1)^2 \to \min_{k_1, k_2}, \tag{23}$$

откуда находим границы $z_3(y) = k_1(y)/n$, $z_4(y) = k_2(y)/n$ допустимой области значений $\hat{F}_2(y)$.

Алгоритм принятия решения. Соотношения (22), (23) используем для построения следующей процедуры проверки статистической гипотезы H_0 о том, что функция $F_0(x)$ является истинной ФРВ исследуемой случайной величины ξ.

1. Для заданного α_0 из условия (23) определяются $k_1(y)$ и $k_2(y)$, затем вычисляются границы $z_3(y)$ и $z_4(y)$ для аргумента $y = j\Delta y$, $j = 1, \ldots, 10n - 1$, $\Delta y = 1/(10n)$.

2. По результатам измерений $x_1, \ldots, x_n$ исследуемой случайной величины ξ находятся величины $F_0(x_1), \ldots, F_0(x_n)$, затем выполняется их сортировка и результат представляется вариационным рядом (20).

3. По формулам (21), (22) вычисляется точечная оценка $\hat{F}_2(y)$ для $y = j\Delta y$, $j = 1, \ldots, 10n - 1$.

4. Решающее правило представляется в виде утверждения: если $\hat{F}_2(j\Delta y) \in [z_3(j\Delta y), z_4(j\Delta y)]$ для всех $j = 1, \ldots, 10n - 1$, то гипотеза H_0 принимается, иначе гипотеза H_0 отклоняется.

Вычисление уровня значимости. Если гипотеза H_0 верна, то в соответствии с условием (23) величины z_3, z_4 не зависят от вида функции $F_0(x)$. При этом $\hat{F}_2$ — точечная оценка ФРВ случайной величины, равномерно распределённой на интервале [0, 1]. Поэтому уровень значимости α предлагаемого критерия согласия не зависит от вида функции $F_0(x)$ и может быть вычислен методом Монте-Карло, для чего используем представленный алгоритм проверки статистической гипотезы со следующими изменениями. На шаге 2 выборка $x_1, \ldots, x_n$ формируется датчиком случайных чисел с равномерным на интервале [0, 1] распределением вероятностей. В результате сортировки по возрастанию значений данной выборки получаем вариационный ряд (20). Остальные шаги алгоритма остаются

Таблица 1

α	α_0			
	$n = 50$	$n = 100$	$n = 500$	$n = 1000$
0,3	$1{,}35 \cdot 10^{-2}$	$1{,}14 \cdot 10^{-2}$	$0{,}79 \cdot 10^{-2}$	$0{,}72 \cdot 10^{-2}$
0,2	$8{,}0 \cdot 10^{-3}$	$6{,}7 \cdot 10^{-3}$	$4{,}6 \cdot 10^{-3}$	$4{,}2 \cdot 10^{-3}$
0,1	$3{,}4 \cdot 10^{-3}$	$2{,}9 \cdot 10^{-3}$	$2{,}1 \cdot 10^{-3}$	$1{,}8 \cdot 10^{-3}$
0,05	$1{,}61 \cdot 10^{-3}$	$1{,}17 \cdot 10^{-3}$	$0{,}92 \cdot 10^{-3}$	$0{,}78 \cdot 10^{-3}$
0,02	$5{,}1 \cdot 10^{-4}$	$4{,}8 \cdot 10^{-4}$	$2{,}9 \cdot 10^{-4}$	$2{,}5 \cdot 10^{-4}$
0,01	$2{,}2 \cdot 10^{-4}$	$1{,}8 \cdot 10^{-4}$	$1{,}5 \cdot 10^{-4}$	$1{,}1 \cdot 10^{-4}$

без изменений. Таким образом, на шаге 4 гипотеза H_0 может быть принята (верное решение) или отклонена (неверное решение). Пусть шаги 2–4 выполнены N раз и гипотеза H_0 отклонена N_0 раз, тогда N_0/N — частота события «отклонить H_0» при условии, что H_0 верна. Известно, что частота сходится в среднеквадратическом к вероятности события, поэтому при достаточно большом N частота N_0/N близка к уровню значимости α.

В табл. 1 представлены результаты вычислений уровня значимости α и соответствующего параметра α_0 для четырёх размеров выборки n равных 50, 100, 500 и 1000. Вычисления выполнялись для $N = 10^4$, что определяет точность полученных результатов до второго знака.

Численное моделирование. Для оценки качества критерия согласия предлагается следующий численный эксперимент. Генерируется m независимых выборок $x_1, \ldots, x_n$ объёма n случайной величины ξ с известной функцией распределения вероятностей $F(x)$. В каждом эксперименте проверяется шесть гипотез H_L, H_0, H_1, H_2, H_3, H_4 о принадлежности выборки соответствующему закону распределения, одна из которых является истинной. Гипотеза H_L соответствует распределению Лапласа с нулевым математическим ожиданием и дисперсией равной 1,65. Данное значение дисперсии выбрано из условия минимального расстояния $\int\limits_{-\infty}^{\infty} |f_L(x) - f_G(x)| dx$ между функциями f_L — плотностью вероятности распределения Лапласа и f_G — плотностью вероятности нормальной случайной величины с нулевым математическим ожиданием и единичной дисперсией. Гипотезы $H_0, \ldots, H_4$ соответствуют значениям параметра β: 1, 1,5, 2, 2,5, 3 семейства распределений с плотностью вероятностей вида [2]

$$f(x) = \frac{\beta}{2\gamma\Gamma(1/\beta)} \exp\Big[-\Big(\frac{|x|}{\gamma}\Big)^{\beta}\Big], \tag{24}$$

где значение параметра $\gamma = \sqrt{\Gamma(1/\beta)/\Gamma(3/\beta)}$ фиксировано в целях нормировки дисперсии на единицу; $\Gamma(x)$ — гамма-функция аргумента x. Семейство (24) включает в себя нормальное распределение (при $\beta = 2$) и распределение Лапласа (при $\beta = 1$).

Рассматривались три варианта истинной гипотезы: H_L, H_2 и H_4. Каждый эксперимент выполнялся для следующих значений уровня значимости α: 0,01, 0,02, 0,05, 0,1, 0,2, 0,3 и объёмов выборки n: 50, 100, 500, 1000. Число выборок m для всех экспериментов равно 1000. Для каждой гипотезы H_i вычислялась частота $\nu_1(H_i)$ как отношение числа выборок, для которых наблюдалось событие «принята гипотеза H_i», к общему числу выборок m. Истинность гипотезы H_i проверялась в каждом эксперименте с использованием критерия согласия Колмогорова, критерия хи-квадрат и предложенного алгоритма. В методе хи-квадрат число интервалов выбиралось равным $\sqrt{n}$. Качество критериев согласия можно

Таблица 2

Критерий	$n = 100$						
	$\nu_1(H_L)$	$\nu_1(H_0)$	$\nu_1(H_1)$	$\nu_1(H_2)$	$\nu_1(H_3)$	$\nu_1(H_4)$	Δ
Колмогорова	0,909	0,702	0,894	0,864	0,817	0,771	0,0994
хи-квадрат	0,902	0,584	0,740	0,686	0,586	0,488	0,2852
предлагаемый	0,881	0,373	0,356	0,201	0,091	0,054	0,670
	$n = 500$						
Колмогорова	0,904	0,113	0,683	0,581	0,282	0,112	0,550
хи-квадрат	0,903	0,038	0,088	0,011	0,001	0,000	0,875
предлагаемый	0,891	0,002	0,000	0,000	0,000	0,000	0,891

оценить величиной $\Delta = \sum\limits_{i:\, H_i \neq H} |\nu_1(H) - \nu_1(H_i)|/5$, где H — истинная гипотеза, и суммирование выполняется по всем альтернативным гипотезам H_i. Значение этой величины лежит в интервале $[0; 1-\alpha]$, причём $1-\alpha$ соответствует идеальному алгоритму с уровнем значимости α, а 0 — случайному принятию решения. Если по условиям рассматриваемой задачи можно выбрать $\alpha \approx 0$, то идеальному алгоритму соответствует значение $\Delta \approx 1$.

Результаты эксперимента показали, что предлагаемый критерий согласия имеет преимущество для выборок небольшого объёма и для истинного ЗРВ с большей дисперсией, чем альтернативные. С ростом объёма выборки рассматриваемые критерии становятся эквивалентными по качеству. В табл. 2 приведены частоты $\nu_1(H_i)$ и показатель качества Δ для значения $\alpha = 0{,}1$, объёмов выборки $n = 100$ и 500 при условии, что верна гипотеза H_L. Для относительно малого объёма выборки $n = 100$ показатель качества Δ критериев Колмогорова, хи-квадрат и предложенного равен соответственно 0,1, 0,28 и 0,67. Таким образом, в этих условиях последний критерий значительно эффективнее: его показатель Δ превышает показатель качества для критерия Колмогорова и хи-квадрат соответственно на величины 0,57 и 0,39. С ростом n рассматриваемые критерии становятся эквивалентными. Так, показатель Δ критериев Колмогорова, хи-квадрат и предложенного равен соответственно 0,55, 0,87, 0,90 для $n = 500$ и 0,81, 0,91, 0,90 для $n = 1000$.

В первом приближении можно полагать, что рассматриваемые критерии согласия эквивалентны по качеству, если верна гипотеза H_2 или H_4. Более детально отметим, что при этом показатель Δ критерия хи-квадрат превышает показатель Δ критерия Колмогорова. Превышение несущественно за исключением выбора уровня значимости α слишком малым. Так, максимальная разность показателей составила величину 0,22 для $\alpha = 0{,}01$ ($\beta = 2$, $n = 500$ и $n = 1000$).

Заключение. В данной работе на основе биномиального распределения построен критерий согласия, который сводится к последовательности двусторонних критериев проверки гипотезы о значении ФРВ при разных значениях её аргумента. Доказано, что каждый элемент этой последовательности является несмещённым локально наиболее мощным критерием. Предложена модификация критерия с использованием преобразования исследуемой случайной величины к величине с равномерным ЗРВ. Построен алгоритм вычисления уровня значимости, свободный от распределения вероятностей. Результаты численного моделирования показали, что для выборок небольшого размера и для истинного ЗРВ с большей дисперсией, чем альтернативные, предложенный критерий существенно эффективней, чем критерии Колмогорова и хи-квадрат. С ростом объёма выборки рассматриваемые критерии становятся эквивалентными по качеству. Преимущества предлагаемого критерия согласия можно объяснить тем, что в нём используется полная мера различия эмпирической и теоретической ФРВ, которая задаётся различием этих функций для каждого аргумента из области определения. В реализации алгоритма аргумент полагается дискретным и такая мера представляет собой вектор (или несколько чисел). В отличие от этого в критериях Колмогорова и хи-квадрат подобная мера — единственное число, что в некоторых случаях может не отражать существенных особенностей эмпирической и теоретической ФРВ и приводить к снижению эффективности критерия. Результаты работы можно использовать для решения задач определения вида ФРВ исследуемой случайной величины.

СПИСОК ЛИТЕРАТУРЫ

1. **Лемешко Б. Ю., Постовалов С. Н.** Применение непараметрических критериев согласия при проверке сложных гипотез // Автометрия. 2001. № 2. С. 88–102.
2. **Клявин И. А., Тырсин А. Н.** Метод подбора наилучшего закона распределения случайной величины по экспериментальным данным // Автометрия. 2013. **49**, № 1. С. 18–25.
3. **Лапко А. В., Лапко В. А.** Непараметрические алгоритмы распознавания образов в задаче проверки статистической гипотезы о тождественности двух законов распределения случайных величин // Автометрия. 2010. **46**, № 6. С. 47–53.
4. **Лемешко Б. Ю., Лемешко С. Б., Постовалов С. Н.** Мощность критерия согласия при близких альтернативах // Измерительная техника. 2007. № 2. С. 22–27.
5. **Салов Г. И.** О мощности одного нового статистического критерия и двухвыборочного критерия Вилкоксона // Автометрия. 2014. **50**, № 1. С. 44–59.
6. **Кулешов Е. Л.** Критерий согласия на основе интервальной оценки // Автометрия. 2016. **52**, № 1. С. 30–36.
7. **Kuleshov E. L., Petrov K. A., Kirillova T. S.** Modeling the goodness-of-fit test based on the interval estimation of the probability distribution function // Proc. of the Intern. Conf. on DOOR 2016 CEUR-WS. Vladivostok, Russia, Sept. 19–23, 2016. Vol. 1623. P. 755-763. URL: http://ceur-ws.org/Vol-1623/paperapp10.pdf (дата обращения: 20.06.2017).
8. **Кокс Д., Хинкли Д.** Теоретическая статистика. М.: Мир, 1978. 560 с.
9. **Тихонов А. Н., Арсенин В. Я.** Методы решения некорректных задач. М.: Наука, 1979. 286 с.

Поступила в редакцию 20 июня 2017 г.