АНАЛИЗ И СИНТЕЗ СИГНАЛОВ И ИЗОБРАЖЕНИЙ

УДК 519.24

О РЕШЕНИИ ПРОБЛЕМ ПРИМЕНЕНИЯ НЕКОТОРЫХ НЕПАРАМЕТРИЧЕСКИХ КРИТЕРИЕВ СОГЛАСИЯ*

Б. Ю. Лемешко, А. А. Горбунова, С. Б. Лемешко, А. П. Рогожников

Новосибирский государственный технический университет,
630073, г. Новосибирск, просп. К. Маркса, 20
E-mail: lemeshko@fpm.ami.nstu.ru

Приведены оценки мощности критериев согласия Купера, Ватсона и трёх критериев Жанга со статистиками Z_A, Z_C, Z_K относительно некоторых пар конкурирующих законов при проверке простых и сложных гипотез. Мощности рассматриваемых критериев даны в сравнении с мощностями критериев Колмогорова, Крамера — Мизеса — Смирнова и Андерсона — Дарлинга. Построены модели распределений статистик и таблицы процентных точек, позволяющие применять критерии согласия Купера и Ватсона при проверке сложных гипотез относительно принадлежности выборок различным параметрическим законам распределения. Предложена методика интерактивного моделирования, позволяющая строить и использовать распределения статистик критериев в ходе решения задачи статистического анализа.

Ключевые слова: непараметрические критерии согласия, критерий Купера, критерий Ватсона, критерии Жанга, мощность критерия, простые и сложные гипотезы.

Введение. Перечень критериев согласия, используемых на практике для решения задач статистического анализа экспериментальных результатов, как правило, ограничен критерием χ^2 Пирсона и двумя–тремя непараметрическими критериями: обычно Колмогорова, Крамера — Мизеса — Смирнова и Андерсона — Дарлинга. Примеры их применения зачастую сопровождаются ошибками, приводящими к некорректным выводам.

Используя критерии согласия, необходимо различать проверку простых и сложных гипотез. Простая проверяемая гипотеза имеет вид H_0: $F(x) = F(x,\theta)$, где $F(x,\theta)$ — известная теоретическая функция распределения вероятностей с известным скалярным или векторным параметром θ. При проверке простых гипотез непараметрические критерии согласия являются «свободными от распределения», т. е. распределения $G(S \mid H_0)$ их статистик при справедливости проверяемой гипотезы не зависят от закона $F(x,\theta)$, с которым проверяется согласие.

При проверке сложных гипотез вида H_0: $F(x) \in \{F(x,\theta),\, \theta \in \Theta\}$, когда оценка $\hat\theta$ скалярного или векторного параметра распределения $F(x,\theta)$ вычисляется по той же самой выборке, непараметрические критерии согласия теряют свойство свободы от распределения. При проверке сложных гипотез условные распределения статистик $G(S \mid H_0)$ зависят от ряда факторов: вида наблюдаемого закона $F(x,\theta)$, соответствующего справедливой проверяемой гипотезе H_0; типа оцениваемого параметра и числа оцениваемых параметров; в некоторых случаях от конкретного значения параметра (например, в случаях семейств гамма- и бета-распределений); метода оценивания параметров. Различия в распределениях

*Работа выполнена при поддержке Министерства образования и науки РФ (проект № 8.1274.2011) и Федеральной целевой программы «Научные и научно-педагогические кадры инновационной России» (соглашение № 14.B37.21.0860).

той же самой статистики при проверке простых и сложных гипотез настолько существенны, что пренебрегать этим фактом ни в коем случае нельзя.

В данной работе мы преследовали две цели: во-первых, хотели привлечь внимание к критериям Купера [1], Ватсона [2, 3] и Жанга [4–7], которые в силу объективных причин и отсутствия обоснованных рекомендаций, к нашему сожалению, практически не используются при анализе данных в приложениях; во-вторых, решить существующие проблемы, связанные с применением этих и других критериев при проверке сложных гипотез, а некоторых из них — и при проверке простых.

Подчеркнём, что имеющиеся до сих пор результаты по классическим критериям Купера и Ватсона и по рассмотренным в [4–7] связаны только с проверкой простых гипотез.

Критерий Купера. В работе [1] предложена расширенная статистика критерия типа Колмогорова для проверки гипотезы о том, что случайная выборка принадлежит закону с непрерывной функцией распределения $F(x,\theta)$. Статистика V_n критерия определяется соотношением

$$V_n = \sup_{-\infty < x < \infty}\{F_n(x) - F(x,\theta)\} - \inf_{-\infty < x < \infty}\{F_n(x) - F(x,\theta)\}$$

(здесь $F_n(x)$ — эмпирическая функция распределения) и используется в виде

$$V_n = D_n^+ + D_n^-, \tag{1}$$

где $D_n^+ = \max\{i/n - F(x_i,\theta)\}$, $D_n^- = \max\{F(x_i,\theta) - (i-1)/n\}$, $i = \overline{1,n}$, n — объём выборки, x_i — здесь и далее элементы вариационного ряда, построенного по выборке (выборка, упорядоченная по возрастанию).

Существенным недостатком критерия со статистикой (1) является сильная зависимость распределения $G(V_n \mid H_0)$ статистики Купера от объёма выборки n. Таблицы процентных точек для случая проверки простых гипотез по критерию со статистикой $\sqrt{n}V_n$ можно найти в [8, 9]. В работе [1] в качестве предельного распределения $G(\sqrt{n}V_n \mid H_0)$ статистики $\sqrt{n}V_n$ дана следующая функция распределения [9]:

$$G(s \mid H_0) = 1 - \sum_{m=1}^{\infty} 2(4m^2s^2 - 1)\mathrm{e}^{-2m^2s^2}. \tag{2}$$

В работе [10] для модификации статистики

$$V = V_n\Big(\sqrt{n} + 0{,}155 + \frac{0{,}24}{\sqrt{n}}\Big), \tag{3}$$

распределение которой в меньшей степени, чем распределение $\sqrt{n}V_n$, зависит от n, приведены процентные точки, представленные в строке 2 табл. 1. Зависимостью распределения статистики (3) от объёма выборки можно пренебречь при $n \geq 20$, так как отклонение реального распределения статистики от предельного незначительно и практически не влияет на результаты статистического вывода.

Можно предложить применять в критерии Купера статистику в виде

$$V_n^{\mathrm{mod}} = \sqrt{n}(D_n^+ + D_n^-) + \frac{1}{3\sqrt{n}}, \tag{4}$$

где идея использования поправки вытекает из выражения для статистики критерия согласия Смирнова [11, с. 81]. Зависимостью распределения статистики (4) от объёма выборки можно практически пренебречь при $n \geq 30$.

Таблица 1

Верхние процентные точки распределений статистик критериев Купера и Ватсона и соответствующие им значения вероятностей вида $P(S > S_\alpha \,|\, H_0)$, вычисленные в соответствии с теоретическими законами и по результатам статистического моделирования

Используемые процентные точки и распределения статистик	α				
	0,15	0,10	0,05	0,025	0,01
Процентные точки статистики Купера [10]	1,537	1,620	1,747	1,862	2,001
Распределение (2)	0,149945	0,099797	0,050075	0,025067	0,009994
Результат моделирования	0,149850	0,099636	0,050060	0,025006	0,009942
Распределение (5)	0,150283	0,100049	0,050030	0,024745	0,009503
Процентные точки статистики Ватсона [13]	0,131	0,152	0,187	0,222	0,267
Распределение (7)	0,150602	0,099526	0,049882	0,024998	0,010283
Результат моделирования	0,150357	0,099479	0,049745	0,024865	0,010305
Распределение (9)	0,149243	0,098704	0,050171	0,025747	0,011149

Статистики (3) и (4) имеют одно и то же предельное распределение. При малых n различие между распределениями статистик (3) и (4) достаточно заметное. Однако при $n \geq 20$ в области принятия решения (при значениях функций распределения статистик $G(V \,|\, H_0) > 0{,}9$ и $G(V_n^{\mathrm{mod}} \,|\, H_0) > 0{,}9$) эти распределения практически совпадают.

В качестве модели предельного закона можно использовать бета-распределение 3-го рода с плотностью

$$f(x) = \frac{\theta_2^{\theta_0}}{\theta_3 B(\theta_0, \theta_1)} \, \frac{((x-\theta_4)/\theta_3)^{\theta_0 - 1}(1-(x-\theta_4)/\theta_3)^{\theta_1 - 1}}{[1+(\theta_2 - 1)((x-\theta_4)/\theta_3)]^{\theta_0 + \theta_1}} \tag{5}$$

($B(\cdot)$ — бета-функция) и вектором параметров $\boldsymbol{\theta} = (7{,}8624,\ 7{,}6629,\ 2{,}6927,\ 2{,}6373,\ 0{,}495)^T$, построенное по результатам моделирования распределения статистики (4). Эта модель хорошо описывает распределение статистики (4) на всей области определения и наряду с предельным может применяться для вычисления достигаемого уровня значимости $P\{S > S^* \,|\, H_0\}$ — вероятности того, что при справедливости проверяемой гипотезы H_0 статистика S критерия превысит величину S^*, где S^* — значение статистики, вычисленное по выборке.

Критерий Ватсона. Статистика критерия Ватсона [2, 3] имеет вид

$$U_n^2 = n \int_{-\infty}^{\infty} \Big\{ F_n(x) - F(x,\theta) - \int_{-\infty}^{\infty} (F_n(y) - F(y,\theta)) dF(y,\theta) \Big\}^2 dF(x,\theta)$$

и используется в следующей удобной для расчётов форме:

$$U_n^2 = \sum_{i=1}^{n} \Big(F(x_i,\theta) - \frac{i-1/2}{n} \Big)^2 - n \Big(\frac{1}{n} \sum_{i=1}^{n} F(x_i,\theta) - \frac{1}{2} \Big)^2 + \frac{1}{12n}. \tag{6}$$

Процентные точки статистики U_n^2 при проверке простой гипотезы можно найти в работах [3, 12]. Предельное распределение $G(U_n^2 \mid H_0)$ статистики U_n^2 приведено в [2, 3] в виде

$$G(s \mid H_0) = 1 - 2 \sum_{m=1}^{\infty} (-1)^{m-1} \mathrm{e}^{-2m^2\pi^2 s}. \tag{7}$$

Модификации критериев Купера и Ватсона рассматривались в [13], критерия Ватсона — в [14]. Асимптотическая эффективность критерия Ватсона исследовалась в [15].

В работе [13] процентные точки приведены для распределений модифицированных статистик. В частности, верхние процентные точки для модифицированной статистики Ватсона в форме

$$U_n^{2*} = (U_n^2 - 0{,}1/n + 0{,}1/n^2)(1 + 0{,}8/n) \tag{8}$$

принимают значения [13], приведённые в строке 6 табл. 1. При объёмах выборок $n \geq 20$ отличием распределения статистики (8) от предельного распределения можно пренебречь.

Практически те же значения верхних процентных точек используют для распределения статистики (6). Следует подчеркнуть, что зависимость распределения статистики (6) от объёма выборки выражена слабо.

Предельное распределение статистики (6) по всей области определения хорошо приближается моделью обратного гауссовского закона с плотностью

$$f(x) = \frac{1}{\theta_2}\Big(\frac{\theta_0}{2\pi((x-\theta_3)/\theta_2)^3}\Big)^{1/2} \exp\Big(-\frac{\theta_0((x-\theta_3)/\theta_2-\theta_1)^2}{2\theta_1^2((x-\theta_3)/\theta_2)}\Big) \tag{9}$$

и вектором параметров $\boldsymbol{\theta} = (0{,}2044,\ 0{,}08344,\ 1{,}0,\ 0{,}0)^T$, построенной по результатам моделирования эмпирического распределения статистики (6). Это распределение при проверке простых гипотез по критерию Ватсона так же, как и предельное, можно использовать для вычисления достигаемого уровня значимости.

В табл. 1 для заданных уровней значимости α представлены значения вероятностей вида $P(S > S_\alpha \mid H_0)$, соответствующие приведённым процентным точкам (критическим значениям) для критерия Купера [10], вычисленные по выражению (2), модели предельного закона (5) и результатам статистического моделирования значений статистики (4) $N = 1{,}7 \cdot 10^6$. В ней же даны аналогичные вероятности для критерия Ватсона [13], вычисленные по соотношению (7), модели предельного закона (9) и результатам статистического моделирования распределения статистики (6). Представленные данные позволяют, с одной стороны, судить о точности моделирования распределений статистик критериев, а с другой — о возможности построения хороших моделей для неизвестных предельных (и непредельных) распределений статистик, позволяющих достаточно точно оценивать по ним достигаемый уровень значимости.

Критерии Жанга. В работах [4–7] предложены непараметрические критерии согласия, статистики которых имеют следующий вид:

$$Z_K = \max_{1 \leq i \leq n} \Big((i - 1/2)\log\Big\{\frac{i-1/2}{nF(x_i,\theta)}\Big\} + (n-i+1/2)\log\Big[\frac{n-i+1/2}{n\{1-F(x_i,\theta)\}}\Big]\Big), \tag{10}$$

$$Z_A = -\sum_{i=1}^{n}\Big[\frac{\log\{F(x_i,\theta)\}}{n-i+1/2} + \frac{\log\{1-F(x_i,\theta)\}}{i-1/2}\Big], \tag{11}$$

$$Z_C = \sum_{i=1}^{n} \Big[\log \Big\{ \frac{[F(x_i, \theta)]^{-1} - 1}{(n-1/2)/(i-3/4) - 1} \Big\} \Big]^2. \tag{12}$$

В справедливости утверждений автора о более высокой мощности критериев по сравнению с критериями Колмогорова, Крамера — Мизеса — Смирнова и Андерсона — Дарлинга мы убедились в [16]. Однако от рекомендации широкого применения критериев со статистиками (10)–(12) остановила сильная зависимость распределений статистик от объёма выборки n. Такая зависимость осложняет использование критериев. Естественно, зависимость от n сохраняется и в случае проверки сложных гипотез.

Анализ мощности критериев. Для исследования распределений статистик при справедливости проверяемой $G(S \mid H_0)$ и конкурирующей $G(S \mid H_1)$ гипотез и вычисления оценок мощности был использован тот же развиваемый подход [17], опирающийся на применение компьютерных технологий и статистическое моделирование. При этом результаты статистического моделирования обеспечивали точность построения распределений статистик $G(S \mid H_i)$, $i = \overline{0,1}$, порядка $\pm 10^{-3}$ с доверительной вероятностью 0,9. Эта величина определяет максимальную длину, которую достигает в районе медианы доверительный интервал, накрывающий истинное значение функции распределения в точке. Для сопоставления мощностей рассматриваемых критериев и других критериев согласия результаты исследований иллюстрируются на двух парах тех же конкурирующих законов, что и в работах [18–20].

Первую пару составили нормальный и логистический законы: проверяемой гипотезе H_0 соответствовал нормальный закон с плотностью

$$f(x) = \frac{1}{\theta_0 \sqrt{2\pi}} \exp\Big\{ -\frac{(x-\theta_1)^2}{2\theta_0^2} \Big\},$$

а конкурирующей гипотезе H_1 — логистический с функцией плотности

$$f(x) = \frac{\pi}{\theta_0\sqrt{3}} \exp\Big\{ -\frac{\pi(x-\theta_1)}{\theta_0\sqrt{3}} \Big\} \Big/ \Big[1 + \exp\Big\{ -\frac{\pi(x-\theta_1)}{\theta_0\sqrt{3}} \Big\} \Big]^2$$

и параметрами $\theta_0 = 1$, $\theta_1 = 0$. В случае простой гипотезы H_0 параметры нормального закона имеют те же значения. Эти два закона близки и трудно различаемы с помощью критериев согласия.

Вторую пару составили: H_0 — распределение Вейбулла с плотностью

$$f(x) = \frac{\theta_0 (x-\theta_2)^{\theta_0 - 1}}{\theta_1^{\theta_0}} \exp\Big\{ -\Big(\frac{x-\theta_2}{\theta_1} \Big)^{\theta_0} \Big\}$$

и параметрами $\theta_0 = 2$, $\theta_1 = 2$, $\theta_2 = 0$; H_1 — гамма-распределение с плотностью

$$f(x) = \frac{1}{\theta_1 \Gamma(\theta_0)} \Big(\frac{x-\theta_2}{\theta_1} \Big)^{\theta_0 - 1} \mathrm{e}^{-(x-\theta_2)/\theta_1}$$

и параметрами $\theta_0 = 3{,}12154$, $\theta_1 = 0{,}557706$, $\theta_2 = 0$, при которых гамма-распределение наиболее близко к закону Вейбулла.

Мощность исследовалась при проверке простых и сложных гипотез H_0 против простой конкурирующей гипотезы H_1 при различных значениях вероятности ошибок 1-го рода α и различных объёмах выборок n.

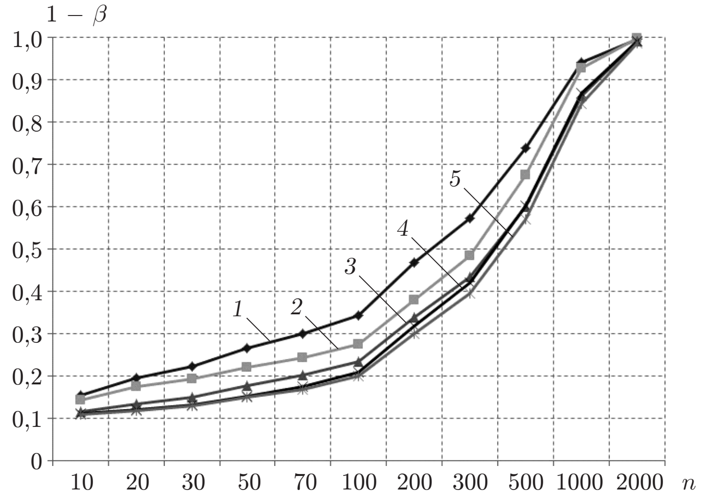

Рис. 1

В качестве примера на рис. 1 для вероятности ошибки 1-го рода $\alpha = 0{,}1$ показаны зависимости мощности критериев от объёмов выборок n при проверке простой гипотезы H_0 (нормальное распределение) против гипотезы H_1 (логистическое).

На этом и последующих рисунках использованы следующие обозначения: кривая *1* — зависимость мощности от объёма выборки для критерия Жанга со статистикой Z_C (12); *2* — для критерия Жанга со статистикой Z_A (11); *3* — для критерия Жанга со статистикой Z_K (10); *4* — для критерия Ватсона со статистикой U_n^2 (6); *5* — для критерия Купера со статистикой V_n^{mod} (5). Аналогично на рис. 2 приведены зависимости для мощности при проверке простой гипотезы H_0 (распределение Вейбулла с параметрами 2, 2, 0) против гипотезы H_1 (гамма-распределение с параметрами 3,12154, 0,557706, 0).

При проверке сложных гипотез в качестве метода оценивания параметров закона использовался метод максимального правдоподобия. Соответствующие зависимости относительно тех же пар конкурирующих законов при проверке сложных гипотез представлены на рис. 3 и 4.

Сравнивая оценки мощности рассматриваемых критериев с результатами для критериев Колмогорова (K), Крамера — Мизеса — Смирнова (KMS) и Андерсона — Дарлинга

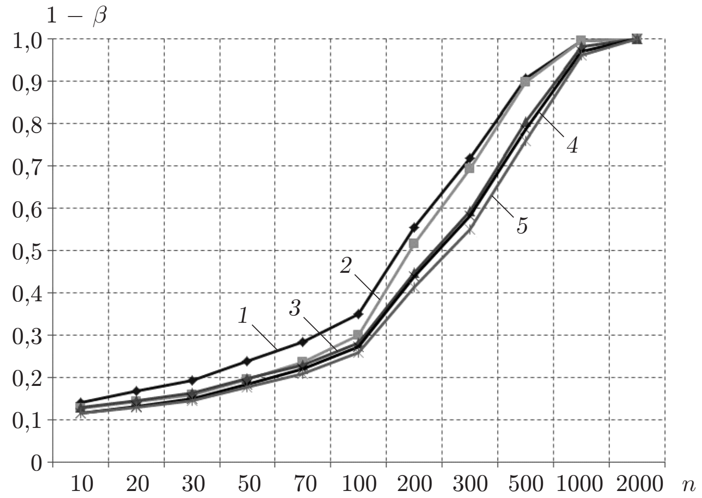

Рис. 2

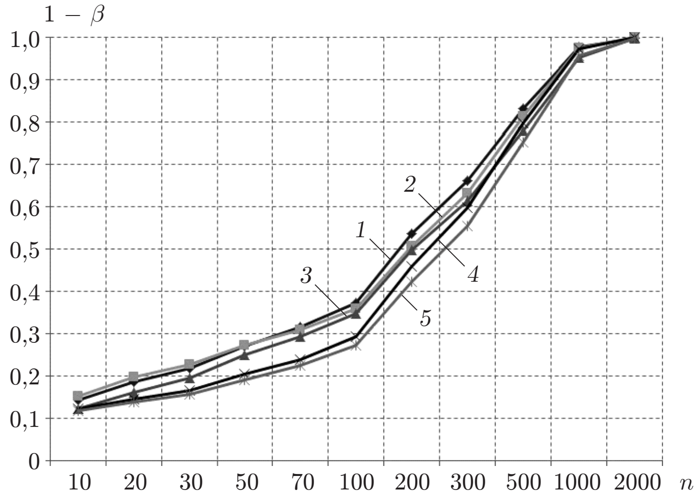

Рис. 3

(AD), приведёнными в работах [19, 20], можно упорядочить критерии по мощности следующим образом:

— при проверке простых гипотез относительно пары «нормальный закон — логистический закон»: $Z_C \succ Z_A \succ Z_K \succ U_n^2 \succ V_n \succ AD \succ K \succ KMS$;

— при проверке простых гипотез относительно пары «закон Вейбулла — гамма-распределение»: $Z_C \succ Z_A \succ Z_K \succ U_n^2 \succ V_n \succ AD \succ KMS \succ K$;

— при проверке сложных гипотез относительно пары «нормальный закон — логистический закон»: $Z_A \approx Z_C \succ Z_K \succ AD \succ KMS \succ U_n^2 \succ V_n \succ K$;

— при проверке сложных гипотез относительно пары «закон Вейбулла — гамма-распределение»: $Z_A \succ Z_C \succ AD \succ Z_K \succ KMS \succ U_n^2 \succ V_n \succ K$.

Если сравнить полученные результаты с мощностью критериев типа χ^2, то при проверке простых гипотез [19] критерий χ^2 Пирсона оказывается на третьей позиции при условиях использования асимптотически оптимального группирования [17] и выбора числа интервалов, при котором критерий будет иметь максимальную мощность [17, 21]. В случае же проверки сложных гипотез [20] позиции критериев χ^2 Пирсона и χ^2 Никулина —

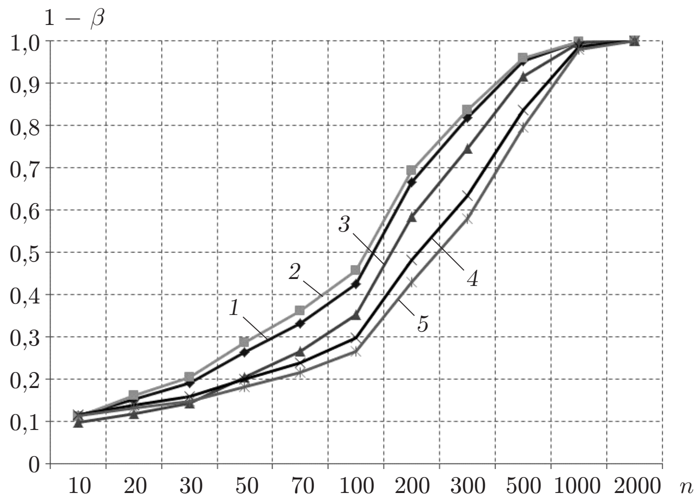

Рис. 4

Рао — Робсона [22–24] ухудшаются: они оказываются на седьмой и восьмой позициях в общем ряду критериев по убыванию мощности. Однако заметим, что мощности этих критериев можно максимизировать относительно заданной конкурирующей гипотезы за счёт оптимального выбора границ и числа интервалов группирования [17, 21].

Применение критериев Купера и Ватсона при проверке сложных гипотез. Как упоминалось выше, при проверке сложных гипотез вида H_0: $F(x) \in \{F(x,\theta),\, \theta \in \Theta\}$, если оценка $\hat{\theta}$ скалярного или векторного параметра распределения $F(x,\theta)$ вычисляется по той же самой выборке, непараметрические критерии согласия теряют свойство свободы от распределения [25].

К решению проблемы проверки сложных гипотез с применением непараметрических критериев согласия использовались различные подходы [10, 13, 26–30]. В работах [31, 32] мы опирались на численные методы исследования и статистическое моделирование. На базе полученных результатов были разработаны рекомендации по применению непараметрических критериев согласия Колмогорова, Крамера — Мизеса — Смирнова, Андерсона — Дарлинга [33, 34]. Позднее эти результаты были уточнены и расширены [35–44] и в настоящее время наиболее полно представлены в [17].

В предлагаемой работе мы приводим таблицы процентных точек и модели распределений статистик критериев Купера и Ватсона, которые рекомендуется использовать при проверке сложных гипотез относительно некоторых часто применяемых в приложениях параметрических законов распределения случайных величин.

Табл. 2 содержит список законов распределения, относительно которых можно проверять сложные гипотезы, используя построенные в данной работе приближения для предельных распределений статистик непараметрических критериев согласия.

Таблицы процентных точек и модели распределений статистик критериев строились по смоделированным выборкам статистик объёмом $N = 1{,}7 \cdot 10^6$. При таких N величина разности между истинным законом $G(S \mid H_0)$ распределения статистики и смоделированным эмпирическим $G_N(S \mid H_0)$ по модулю не превышает величины 10^{-3}. Значения статистик критериев вычислялись по выборкам псевдослучайных величин объёмом $n = 10^3$, генерируемым в соответствии с наблюдаемым законом $F(x,\theta)$. В такой ситуации распределение $G(S_n \mid H_0)$ практически совпадает с предельным $G(S \mid H_0)$. В задачах статистического анализа можно пользоваться представленными в данной работе моделями, начиная с объёмов выборок $n > 25$.

Распределения $G(S \mid H_0)$ статистик критериев Купера и Ватсона наилучшим образом аппроксимируются семейством бета-распределений 3-го рода $B_3(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$, функция плотности которого определяется соотношением (5), и семейством распределений Sl-Джонсона ($Sl(\theta_0, \theta_1, \theta_2, \theta_3)$) (см. табл. 2).

Верхние процентные точки и построенные модели для предельных распределений статистики критерия Купера в случае использования оценок максимального правдоподобия (ОМП) представлены в табл. 3 для законов: экспоненциального, полунормального, Рэлея, Максвелла, Лапласа, нормального, логнормального, Коши, логистического, экстремальных значений (maximum и minimum), Вейбулла. Для тех же законов распределения верхние процентные точки и построенные модели для распределений статистики критерия Ватсона приведены в табл. 4.

В табл. 5 представлены верхние процентные точки и модели предельных распределений статистик в случае проверки сложных гипотез относительно законов распределения Sb-Джонсона (при использовании ОМП), в табл. 6 — относительно законов Sl-Джонсона, в табл. 7 — относительно законов Su-Джонсона.

Точность построенных моделей распределений статистик при проверке сложных гипотез и возможность использования этих моделей в задачах статистического анализа в определённой степени можно продемонстрировать следующим образом.

Таблица 2

Законы распределения случайных величин

Наименование закона	Функция плотности $f(x,\theta)$	Наименование закона	Функция плотности $f(x,\theta)$
Экспоненциальный	$\frac{1}{\theta_0}\,\mathrm{e}^{-x/\theta_0}$	Лапласа	$\frac{1}{2\theta_0}\,\mathrm{e}^{-\lvert x-\theta_1\rvert/\theta_0}$
Полунормальный	$\frac{2}{\theta_0\sqrt{2\pi}}\,\mathrm{e}^{-x^2/2\theta_0^2}$	Нормальный (Гаусса)	$\frac{1}{\theta_0\sqrt{2\pi}}\,\mathrm{e}^{-(x-\theta_1)^2/2\theta_0^2}$
Рэлея	$\frac{x}{\theta_0^2}\,\mathrm{e}^{-x^2/2\theta_0^2}$	Логарифмически нормальный	$\frac{1}{x\theta_0\sqrt{2\pi}}\,\mathrm{e}^{-(\ln x-\theta_1)^2/2\theta_0^2}$
Максвелла	$\frac{2x^2}{\theta_0^3\sqrt{2\pi}}\,\mathrm{e}^{-x^2/2\theta_0^2}$	Коши	$\frac{\theta_0}{\pi[\theta_0^2+(x-\theta_1)^2]}$

Наименование закона	Функция плотности $f(x,\theta)$
Логистический	$\frac{\pi}{\theta_0\sqrt{3}}\exp\Big\{-\frac{\pi(x-\theta_1)}{\theta_0\sqrt{3}}\Big\}\Big/\Big[1+\exp\Big\{-\frac{\pi(x-\theta_1)}{\theta_0\sqrt{3}}\Big\}\Big]^2$
Экстремального значения (maximum)	$\frac{1}{\theta_0}\exp\Big\{-\frac{x-\theta_1}{\theta_0}-\exp\Big(-\frac{x-\theta_1}{\theta_0}\Big)\Big\}$
Экстремального значения (minimum)	$\frac{1}{\theta_0}\exp\Big\{\frac{x-\theta_1}{\theta_0}-\exp\Big(\frac{x-\theta_1}{\theta_0}\Big)\Big\}$
Вейбулла	$\frac{\theta_0 x^{\theta_0-1}}{\theta_1^{\theta_0}}\exp\{-(x/\theta_1)^{\theta_0}\}$
Sb-Джонсона ($Sb(\theta_0,\theta_1,\theta_2,\theta_3)$)	$\frac{\theta_1\theta_2}{(x-\theta_3)(\theta_2+\theta_3-x)}\exp\Big\{-\frac{1}{2}\Big[\theta_0-\theta_1\ln\frac{x-\theta_3}{\theta_2+\theta_3-x}\Big]^2\Big\}$
Sl-Джонсона ($Sl(\theta_0,\theta_1,\theta_2,\theta_3)$)	$\frac{\theta_1}{(x-\theta_3)\sqrt{2\pi}}\exp\Big\{-\frac{1}{2}\Big[\theta_0+\theta_1\ln\frac{x-\theta_3}{\theta_2}\Big]^2\Big\}$
Su-Джонсона ($Su(\theta_0,\theta_1,\theta_2,\theta_3)$)	$\frac{\theta_1}{\sqrt{2\pi}\sqrt{(x-\theta_3)^2+\theta_2^2}}\exp\Big\{-\frac{1}{2}\Big[\theta_0+\theta_1\ln\Big\{\frac{x-\theta_3}{\theta_2}+\sqrt{\Big(\frac{x-\theta_3}{\theta_2}\Big)^2+1}\Big\}\Big]^2\Big\}$

Табл. 8 иллюстрирует результаты построения моделей распределений статистик для проверки сложных гипотез о согласии анализируемой выборки с логистическим законом с вычислением по выборке ОМП параметров сдвига и масштаба этого закона. Построенные модели приведены для критерия Купера в табл. 3, для критерия Ватсона в табл. 4.

В колонке $F_N^1(v)$ табл. 8 представлены оценки функции распределения статистики Купера, соответствующие значениям статистики v и полученные по смоделированному эмпирическому распределению статистики, по которому была построена модель $B_3(14{,}3460, 18{,}6137, 3{,}6366, 3{,}9560, 0{,}3525)$ из табл. 3. Соответствующие значения функции распределения статистики, вычисленные по этой модели, даны в колонке $F(v)$. Колонка $F_N^2(v)$ содержит оценки функции распределения, полученные по контрольной (вновь сгенерированной) выборке статистик.

Аналогично соответствующие оценки для функции распределения статистики критерия Ватсона при построении модели $B_3(4{,}2608, 4{,}6784, 9{,}3054, 0{,}3810, 0{,}0084)$ из табл. 4 приведены в колонках $F_N^1(u_n^2)$, $F(u_n^2)$, $F_N^2(u_n^2)$ табл. 8.

Таблица 3

Верхние процентные точки и модели предельных распределений статистики критерия Купера в случае использования ОМП

Наименование закона	Оцениваемый параметр	Процентные точки			Модель
		0,1	0,05	0,01	
Экспоненциальный, Рэлея, Максвелла	Масштаб	1,540	1,661	1,905	$B_3(5{,}5932, 7{,}6149, 2{,}1484, 2{,}3961, 0{,}5630)$
Полунормальный	Масштаб	1,543	1,664	1,907	$B_3(11{,}4707, 40{,}7237, 7{,}020, 20{,}3675, 0{,}3989)$
Лапласа	Масштаб	1,469	1,587	1,825	$B_3(7{,}8324, 8{,}3778, 2{,}6906, 2{,}4820, 0{,}4830)$
	Сдвиг	1,473	1,597	1,850	$B_3(9{,}1630, 6{,}6097, 4{,}0210, 2{,}4081, 0{,}4900)$
	Оба параметра	1,278	1,365	1,541	$B_3(10{,}0376, 7{,}8452, 3{,}4694, 1{,}9586, 0{,}4756)$
Нормальный, логарифмически нормальный	Масштаб	1,494	1,611	1,847	$B_3(6{,}3057, 8{,}1797, 2{,}3279, 2{,}4413, 0{,}5370)$
	Сдвиг	1,540	1,662	1,908	$B_3(5{,}5932, 7{,}6149, 2{,}1484, 2{,}3961, 0{,}5630)$
	Оба параметра	1,402	1,505	1,709	$B_3(7{,}4917, 8{,}0016, 2{,}4595, 2{,}1431, 0{,}4937)$
Коши	Масштаб или сдвиг	1,435	1,560	1,815	$B_3(3{,}8425, 5{,}9345, 2{,}4284, 2{,}1927, 0{,}6150)$
	Оба параметра	1,126	1,197	1,337	$B_3(9{,}4267, 7{,}5349, 3{,}2515, 1{,}5491, 0{,}4700)$
Логистический	Масштаб	1,470	1,588	1,826	$B_3(9{,}7224, 7{,}8186, 3{,}2399, 2{,}4541, 0{,}4370)$
	Сдвиг	1,511	1,633	1,880	$B_3(9{,}1363, 6{,}9693, 3{,}4630, 2{,}3985, 0{,}4790)$
	Оба параметра	1,337	1,432	1,622	$B_3(14{,}3460, 18{,}6137, 3{,}6366, 3{,}9560, 0{,}3525)$
Экстремальных значений, Вейбулла	Масштаб*	1,504	1,622	1,861	$Sl(1{,}2459, 4{,}0123, 1{,}3063, 0{,}1873)$
	Сдвиг**	1,540	1,662	1,908	$B_3(5{,}5932, 7{,}6149, 2{,}1484, 2{,}3961, 0{,}5630)$
	Оба параметра	1,411	1,516	1,726	$Sl(1{,}4012, 5{,}0846, 1{,}4465, -0{,}0070)$

Примечания: $*$ — при оценивании параметра формы распределения Вейбулла, $**$ — при оценивании параметра масштаба распределения Вейбулла.

Как видно из табл. 8 для критерия Купера $\max|F_N^1(v) - F_N^2(v)| = 0{,}000522$, что укладывается в 90 %-ный доверительный интервал, максимальная величина которого при $N = 1{,}7\cdot 10^6$ не превышает 0,002. Отклонения эмпирических распределений от построенной модели $\max|F_N^1(v) - F(v)| = 0{,}001509$, $\max|F(v) - F_N^2(v)| = 0{,}001579$. Аналогично для критерия Ватсона имеем $\max|F_N^1(u_n^2) - F_N^2(u_n^2)| = 0{,}000537$, $\max|F_N^1(u_n^2) - F(u_n^2)| = 0{,}000932$, $\max|F(u_n^2) - F_N^2(u_n^2)| = 0{,}001323$.

Интерактивный режим исследования распределений статистик. Самым серьёзным препятствием, возникающим на пути решения проблемы применения непараметрических критериев согласия для проверки сложных гипотез относительно широкого круга возможных параметрических законов распределения, используемых в различных приложениях для описания наблюдаемых случайных величин (ошибок измерения), является зависимость распределений статистик критериев от конкретных значений параметра (или параметров) формы закона, соответствующего проверяемой гипотезе (в случае семейств гамма- и бета-распределений, обобщённого Вейбулла, обратного гауссовского и др.). Как правило, это касается законов, применение которых наиболее перспективно в различных приложениях, в задачах анализа выживания и исследования надёжности сложных изделий и систем.

Таблица 4

Верхние процентные точки и модели предельных распределений статистики критерия Ватсона в случае использования ОМП

Наименование закона	Оцениваемый параметр	Процентные точки			Модель
		0,1	0,05	0,01	
Экспоненциальный, Рэлея, Максвелла	Масштаб	0,129	0,159	0,230	$B_3(4{,}0419, 2{,}9119, 10{,}5931, 0{,}5000, 0{,}0096)$
Полунормальный	Масштаб	0,131	0,161	0,232	$B_3(4{,}9988, 3{,}8721, 15{,}1781, 0{,}6900, 0{,}0059)$
Лапласа	Масштаб	0,115	0,144	0,214	$B_3(9{,}2136, 3{,}8610, 30{,}5491, 0{,}7010, 0{,}0015)$
	Сдвиг	0,111	0,139	0,209	$B_3(7{,}4479, 3{,}2650, 30{,}7784, 0{,}6227, 0{,}0063)$
	Оба параметра	0,071	0,084	0,114	$B_3(9{,}0116, 5{,}3554, 17{,}3201, 0{,}3908, 0{,}0038)$
Нормальный, логарифмически нормальный	Масштаб	0,122	0,151	0,221	$B_3(8{,}8122, 3{,}7536, 29{,}8074, 0{,}7171, 0{,}0019)$
	Сдвиг	0,127	0,157	0,228	$B_3(3{,}6769, 4{,}4438, 9{,}8994, 0{,}6805, 0{,}0082)$
	Оба параметра	0,096	0,116	0,164	$B_3(3{,}5230, 4{,}4077, 9{,}2281, 0{,}4785, 0{,}0104)$
Коши	Масштаб или сдвиг	0,105	0,133	0,203	$Sl(2{,}7778, 1{,}5065, 0{,}2690, 0{,}0049)$
	Оба параметра	0,052	0,061	0,081	$B_3(8{,}3558, 4{,}8650, 12{,}0768, 0{,}1930, 0{,}0049)$
Логистический	Масштаб	0,115	0,144	0,214	$B_3(9{,}2136, 3{,}8610, 30{,}5491, 0{,}7010, 0{,}0015)$
	Сдвиг	0,119	0,148	0,218	$B_3(3{,}9730, 3{,}9414, 13{,}2655, 0{,}6637, 0{,}0090)$
	Оба параметра	0,081	0,098	0,135	$B_3(4{,}2608, 4{,}6784, 9{,}3054, 0{,}3810, 0{,}0084)$
Экстремальных значений, Вейбулла	Масштаб*	0,122	0,151	0,221	$B_3(8{,}8122, 3{,}7536, 29{,}8074, 0{,}7171, 0{,}0019)$
	Сдвиг**	0,129	0,159	0,230	$B_3(4{,}9988, 3{,}8721, 15{,}1781, 0{,}6792, 0{,}0061)$
	Оба параметра	0,097	0,118	0,165	$Sl(1{,}2863, 1{,}6736, 0{,}0927, 0{,}0052)$

Примечания: $*$ — при оценивании параметра формы распределения Вейбулла, $**$ — при оценивании параметра масштаба распределения Вейбулла.

Таблица 5

Верхние процентные точки и модели предельных распределений статистик непараметрических критериев согласия в случае проверки гипотез относительно распределений Sb-Джонсона при использовании ОМП

Оцениваемый параметр	Процентные точки			Модель
	0,1	0,05	0,01	
для критерия Купера				
θ_0	1,540	1,662	1,908	$B_3(5{,}5932, 7{,}6149, 2{,}1484, 2{,}3961, 0{,}5630)$
θ_1	1,494	1,611	1,847	$B_3(6{,}3057, 8{,}1797, 2{,}3279, 2{,}4413, 0{,}5370)$
θ_0, θ_1	1,402	1,505	1,709	$B_3(7{,}4917, 8{,}0016, 2{,}4595, 2{,}1431, 0{,}4937)$
для критерия Ватсона				
θ_0	0,127	0,157	0,228	$B_3(3{,}6769, 4{,}4438, 9{,}8994, 0{,}6805, 0{,}0082)$
θ_1	0,122	0,151	0,221	$B_3(8{,}8122, 3{,}7536, 29{,}8074, 0{,}7171, 0{,}0019)$
θ_0, θ_1	0,096	0,116	0,164	$B_3(3{,}5230, 4{,}4077, 9{,}2281, 0{,}4785, 0{,}0104)$

Таблица 6

Верхние процентные точки и модели предельных распределений статистик непараметрических критериев согласия в случае проверки гипотез относительно распределений Sl-Джонсона при использовании ОМП

Оцениваемый параметр	Процентные точки			Модель
	0,1	0,05	0,01	
для критерия Купера				
θ_0	1,540	1,662	1,908	$B_3(5{,}5932, 7{,}6149, 2{,}1484, 2{,}3961, 0{,}5630)$
θ_1	1,512	1,631	1,872	$B_3(6{,}7423, 8{,}0549, 2{,}4935, 2{,}4976, 0{,}5250)$
θ_2	1,540	1,662	1,908	$B_3(5{,}5932, 7{,}6149, 2{,}1484, 2{,}3961, 0{,}5630)$
θ_0, θ_1	1,402	1,505	1,709	$B_3(7{,}4917, 8{,}0016, 2{,}4595, 2{,}1431, 0{,}4937)$
θ_0, θ_2	1,540	1,662	1,908	$B_3(5{,}5932, 7{,}6149, 2{,}1484, 2{,}3961, 0{,}5630)$
θ_1, θ_2	1,402	1,505	1,709	$B_3(7{,}4917, 8{,}0016, 2{,}4595, 2{,}1431, 0{,}4937)$
$\theta_0, \theta_1, \theta_2$	1,402	1,505	1,709	$B_3(7{,}4917, 8{,}0016, 2{,}4595, 2{,}1431, 0{,}4937)$
для критерия Ватсона				
θ_0	0,127	0,157	0,228	$B_3(3{,}6769, 4{,}4438, 9{,}8994, 0{,}6805, 0{,}0082)$
θ_1	0,124	0,153	0,223	$B_3(3{,}4122, 4{,}9262, 9{,}6902, 0{,}7643, 0{,}0087)$
θ_2	0,127	0,157	0,228	$B_3(3{,}6769, 4{,}4438, 9{,}8994, 0{,}6805, 0{,}0082)$
θ_0, θ_1	0,096	0,116	0,164	$B_3(3{,}5230, 4{,}4077, 9{,}2281, 0{,}4785, 0{,}0104)$
θ_0, θ_2	0,127	0,157	0,228	$B_3(3{,}6769, 4{,}4438, 9{,}8994, 0{,}6805, 0{,}0082)$
θ_1, θ_2	0,096	0,116	0,164	$B_3(3{,}5230, 4{,}4077, 9{,}2281, 0{,}4785, 0{,}0104)$
$\theta_0, \theta_1, \theta_2$	0,096	0,116	0,164	$B_3(3{,}5230, 4{,}4077, 9{,}2281, 0{,}4785, 0{,}0104)$

Так как оценки параметров становятся известными только в процессе анализа, требуемое для проверки гипотезы распределение статистики нельзя найти заранее (до вычисления оценок по анализируемой выборке!). В случае критериев со статистиками (10)–(12) проблема усугубляется зависимостью распределений статистик от объёмов выборок. Отсюда следует, что распределения статистик применяемых критериев должны находиться в интерактивном режиме в ходе проводимого статистического анализа [45], а затем использоваться при формировании вывода по итогам проверки сложной гипотезы.

Реализация такого интерактивного режима требует наличия развитого программного обеспечения, позволяющего (как в [46]) в целях ускорения распараллеливать процессы моделирования и привлекать доступные вычислительные ресурсы. В условиях распараллеливания время построения распределения $G_N(S_n \mid H_0)$ статистики критерия (с требуемой точностью), необходимого для проверки гипотезы, и определения по нему достигнутого уровня значимости $P\{S_n \geq S^*\}$, где S^* — вычисленное по выборке значение статистики, оказывается не очень заметным на фоне полного решения задачи статистического анализа.

В работе [46] интерактивный режим исследования распределений статистик реализован для следующих непараметрических критериев согласия: Колмогорова, Крамера — Мизеса — Смирнова, Андерсона — Дарлинга, Купера, Ватсона, Жанга (три критерия). При этом могут использоваться различные методы оценивания параметров.

В задачах анализа выживания и исследования надёжности сложных систем выборки, как правило, характеризуются наличием цензурированных и случайно (многократно) цензурированных наблюдений. Для проверки адекватности построенных функций надёжности также применяются критерии согласия, но уже к формируемым выборкам остатков. В случае цензурированных выборок распределения статистик используемых модифициро-

Таблица 7

Верхние процентные точки и модели предельных распределений статистик непараметрических критериев согласия в случае проверки гипотез относительно распределений Su-Джонсона при использовании ОМП

Оцениваемый параметр	Процентные точки			Модель
	0,1	0,05	0,01	
для критерия Купера				
θ_0	1,540	1,662	1,908	$B_3(5{,}5932, 7{,}6149, 2{,}1484, 2{,}3961, 0{,}5630)$
θ_1	1,512	1,631	1,872	$B_3(6{,}7676, 8{,}3605, 2{,}3501, 2{,}4976, 0{,}5142)$
θ_2	1,491	1,612	1,857	$B_3(7{,}5884, 8{,}1397, 2{,}6781, 2{,}4982, 0{,}4882)$
θ_3	1,517	1,638	1,885	$B_3(8{,}1449, 7{,}2651, 3{,}0338, 2{,}4418, 0{,}4880)$
θ_0, θ_1	1,402	1,505	1,709	$B_3(8{,}1449, 7{,}2650, 3{,}0338, 2{,}1431, 0{,}5015)$
θ_0, θ_2	1,393	1,496	1,703	$B_3(7{,}5234, 7{,}3134, 2{,}7694, 2{,}1076, 0{,}5035)$
θ_0, θ_3	1,390	1,496	1,713	$B_3(8{,}0187, 7{,}7542, 2{,}7862, 2{,}1751, 0{,}4800)$
θ_1, θ_2	1,414	1,525	1,749	$B_3(8{,}6702, 7{,}5387, 2{,}9284, 2{,}2036, 0{,}4600)$
θ_1, θ_3	1,375	1,475	1,675	$B_3(8{,}6702, 7{,}5387, 2{,}9284, 2{,}0887, 0{,}4740)$
θ_2, θ_3	1,350	1,447	1,640	$B_3(9{,}0132, 7{,}9999, 2{,}8585, 2{,}0644, 0{,}4635)$
$\theta_0, \theta_1, \theta_2$	1,324	1,422	1,621	$B_3(10{,}7806, 8{,}4043, 3{,}2432, 2{,}1461, 0{,}4150)$
$\theta_0, \theta_1, \theta_3$	1,333	1,431	1,629	$B_3(10{,}3455, 8{,}0495, 3{,}5687, 2{,}1993, 0{,}4463)$
$\theta_0, \theta_2, \theta_3$	1,296	1,388	1,575	$B_3(10{,}3223, 7{,}7893, 3{,}3393, 2{,}0021, 0{,}4358)$
$\theta_1, \theta_2, \theta_3$	1,299	1,394	1,584	$B_3(10{,}5957, 8{,}2600, 3{,}2334, 2{,}0676, 0{,}4194)$
$\theta_0, \theta_1, \theta_2, \theta_3$	1,235	1,321	1,494	$B_3(9{,}9689, 7{,}3418, 3{,}4037, 1{,}8225, 0{,}4438)$
для критерия Ватсона				
θ_0	0,127	0,157	0,228	$B_3(3{,}6769, 4{,}4438, 9{,}8994, 0{,}6805, 0{,}0082)$
θ_1	0,124	0,153	0,223	$B_3(3{,}4122, 4{,}9262, 9{,}6902, 0{,}7643, 0{,}0087)$
θ_2	0,117	0,146	0,215	$B_3(6{,}0296, 3{,}7175, 22{,}6978, 0{,}7115, 0{,}0057)$
θ_3	0,121	0,150	0,220	$B_3(7{,}4154, 3{,}9208, 22{,}4649, 0{,}6800, 0{,}0022)$
θ_0, θ_1	0,096	0,116	0,164	$B_3(3{,}5230, 4{,}4077, 9{,}2281, 0{,}4785, 0{,}0104)$
θ_0, θ_2	0,093	0,114	0,161	$B_3(4{,}0651, 4{,}8643, 9{,}5614, 0{,}4903, 0{,}0078)$
θ_0, θ_3	0,092	0,113	0,162	$B_3(4{,}4170, 4{,}9456, 10{,}4292, 0{,}5005, 0{,}0067)$
θ_1, θ_2	0,099	0,123	0,181	$B_3(5{,}5181, 4{,}1815, 16{,}0852, 0{,}5478, 0{,}0055)$
θ_1, θ_3	0,089	0,108	0,151	$B_3(5{,}7461, 4{,}4051, 13{,}9768, 0{,}4528, 0{,}0060)$
θ_2, θ_3	0,084	0,101	0,141	$B_3(5{,}9952, 4{,}3409, 13{,}8757, 0{,}4020, 0{,}0060)$
$\theta_0, \theta_1, \theta_2$	0,077	0,093	0,131	$B_3(5{,}5809, 4{,}9570, 14{,}1052, 0{,}4540, 0{,}0060)$
$\theta_0, \theta_1, \theta_3$	0,080	0,097	0,137	$B_3(5{,}8959, 4{,}4478, 14{,}5923, 0{,}4132, 0{,}0060)$
$\theta_0, \theta_2, \theta_3$	0,072	0,087	0,121	$B_3(6{,}1780, 4{,}6712, 14{,}5568, 0{,}3791, 0{,}0060)$
$\theta_1, \theta_2, \theta_3$	0,072	0,087	0,121	$B_3(6{,}1780, 4{,}6712, 14{,}5568, 0{,}3791, 0{,}0060)$
$\theta_0, \theta_1, \theta_2, \theta_3$	0,062	0,074	0,101	$B_3(7{,}3816, 4{,}4215, 14{,}1896, 0{,}2616, 0{,}0053)$

Таблица 8

Функции распределения статистик критериев Купера и Ватсона при проверке согласия с логистическим законом для случая вычисления ОМП двух параметров закона

Для критерия Купера				Для критерия Ватсона			
v	$F_N^1(v)$	$F(v)$	$F_N^2(v)$	u_n^2	$F_N^1(u_n^2)$	$F(u_n^2)$	$F_N^2(u_n^2)$
0,6	0,000869	0,000987	0,000848	0,012	0,001923	0,001182	0,001945
0,7	0,015031	0,015772	0,015026	0,024	0,120052	0,120838	0,119515
0,8	0,081186	0,081282	0,080664	0,036	0,376606	0,375904	0,376442
0,9	0,224404	0,223060	0,224009	0,048	0,602155	0,601577	0,601897
1,0	0,417929	0,417090	0,417482	0,060	0,755732	0,755849	0,755466
1,1	0,610865	0,611452	0,610709	0,072	0,851977	0,852258	0,851492
1,2	0,765689	0,767198	0,765619	0,084	0,910636	0,910653	0,910282
1,3	0,872154	0,872875	0,871804	0,096	0,946156	0,945746	0,946005
1,4	0,936271	0,935934	0,935922	0,108	0,967708	0,966874	0,967672
1,5	0,970832	0,969894	0,970669	0,120	0,980601	0,979669	0,980678
1,6	0,987655	0,986702	0,987742	0,132	0,988400	0,987471	0,988269
1,7	0,995212	0,994446	0,995231	0,144	0,993021	0,992259	0,993041
1,8	0,998289	0,997797	0,998248	0,156	0,995797	0,995213	0,995842
1,9	0,999419	0,999168	0,999416	0,168	0,997516	0,997044	0,997457
2,0	0,999818	0,999701	0,999798	0,180	0,998483	0,998181	0,998442

ванных непараметрических критериев согласия дополнительно зависят от законов распределения моментов цензурирования и степени цензурирования выборок [47–49]. В таких условиях реализация интерактивного режима исследования особенно необходима.

Приводимый далее пример демонстрирует точность определения достигнутого уровня значимости в зависимости от величины выборки N моделируемого в интерактивном режиме эмпирического распределения статистики [46].

П р и м е р. Необходимо проверить сложную гипотезу о принадлежности следующей выборки объёмом $n = 100$ обратному гауссовскому закону с плотностью (9):

0,945	1,040	0,239	0,382	0,398	0,946	1,248	1,437	0,286	0,987
2,009	0,319	0,498	0,694	0,340	1,289	0,316	1,839	0,432	0,705
0,371	0,668	0,421	1,267	0,466	0,311	0,466	0,967	1,031	0,477
0,322	1,656	1,745	0,786	0,253	1,260	0,145	3,032	0,329	0,645
0,374	0,236	2,081	1,198	0,692	0,599	0,811	0,274	1,311	0,534
1,048	1,411	1,052	1,051	4,682	0,111	1,201	0,375	0,373	3,694
0,426	0,675	3,150	0,424	1,422	3,058	1,579	0,436	1,167	0,445
0,463	0,759	1,598	2,270	0,884	0,448	0,858	0,310	0,431	0,919
0,796	0,415	0,143	0,805	0,827	0,161	8,028	0,149	2,396	2,514
1,027	0,775	0,240	2,745	0,885	0,672	0,810	0,144	0,125	1,621

Параметр сдвига $\theta_3 = 0$ предполагается заданным.

По выборке оцениваются параметры формы θ_0, θ_1 и масштабный параметр θ_2. Найденные по данной выборке ОМП параметров: $\hat{\theta}_0 = 0{,}7481$, $\hat{\theta}_1 = 0{,}7808$, $\hat{\theta}_2 = 1{,}3202$. Распределения статистик всех непараметрических критериев согласия в этом случае зависят от

Таблица 9

Достигнутые уровни значимости по критериям согласия при различных N

Значения статистик критериев	$N = 10^3$	$N = 10^4$	$N = 10^5$	$N = 10^6$
$V_n^{\text{mod}} = 1{,}1113$	0,479	0,492	0,493	0,492
$U_n^2 = 0{,}05200$	0,467	0,479	0,483	0,482
$Z_A = 3{,}3043$	0,661	0,681	0,679	0,678
$Z_C = 4{,}7975$	0,751	0,776	0,777	0,776
$Z_K = 1{,}4164$	0,263	0,278	0,272	0,270
$K = 0{,}5919$	0,643	0,659	0,662	0,662
$KMS = 0{,}05387$	0,540	0,557	0,560	0,561
$AD = 0{,}3514$	0,529	0,549	0,548	0,547

значений параметров формы θ_0 и θ_1 [40–42], не зависят от значения параметра масштаба θ_2 и должны быть найдены при значениях $\theta_0 = 0{,}7481$, $\theta_1 = 0{,}7808$.

Вычисленные по выборке значения S_i^* статистик критериев Купера, Ватсона, Жанга, Колмогорова, Крамера — Мизеса — Смирнова, Андерсона — Дарлинга и соответствующие этим значениям достигнутые уровни значимости $P\{S \geq S_i^* \mid H_0\}$ (p-value), полученные при различной точности моделирования (при различном объёме N моделируемых выборок статистик), приведены в табл. 9.

Заключение. В представленной работе показано, что при проверке простых гипотез критерии Купера и Ватсона имеют преимущество в мощности перед критериями Колмогорова, Крамера — Мизеса — Смирнова и Андерсона — Дарлинга. Проблем применения этих критериев при проверке простых гипотез нет.

При проверке сложных гипотез критерии Купера и Ватсона утрачивают имевшееся преимущество в мощности. Однако отсюда не следует отказ от использования данных критериев, так как применение множества критериев, использующих различные меры отклонения эмпирического распределения от теоретического, повышает качество статистических выводов. Построенные в данной работе модели распределений статистик и таблицы процентных точек позволяют корректно применять критерии Купера и Ватсона при проверке сложных гипотез относительно ряда параметрических моделей законов распределения (в случае использования ОМП).

Критерии Жанга со статистиками Z_C и Z_A имеют неоспоримое преимущество в мощности перед всеми остальными. Особенно это заметно при проверке простых гипотез. Определённые трудности при использовании критериев, связанные с существенной зависимостью распределений статистик от объёма выборок, разрешаются за счёт интерактивного режима.

Реализованный интерактивный режим исследования распределений статистик [46] даёт возможность корректного использования критериев согласия Колмогорова, Крамера — Мизеса — Смирнова, Андерсона — Дарлинга, Купера, Ватсона, Жанга (со статистиками Z_C, Z_A, Z_K) и в тех случаях, когда к моменту проверки сложной гипотезы H_0 распределение статистики применяемого критерия, соответствующее справедливости H_0, неизвестно. Для критериев Жанга этот режим обеспечивает проверку и простых гипотез при произвольных объёмах выборок.

СПИСОК ЛИТЕРАТУРЫ

1. **Kuiper N. H.** Tests concerning random points on a circle // Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen. 1960. Ser. A. Vol. 63. P. 38–47.

2. **Watson G. S.** Goodness-of-fit tests on a circle. I // Biometrika. 1961. **48**, N 1–2. P. 109–114.
3. **Watson G. S.** Goodness-of-fit tests on a circle. II // Biometrika. 1962. **49**, N 1–2. P. 57–63.
4. **Zhang J.** Powerful goodness-of-fit and multi-sample tests: PhD Thesis. York University, Toronto. 2001. 113 p. URL: http://www.collectionscanada.gc.ca/obj/s4/f2/dsk3/ftp05/NQ66371.pdf (дата обращения: 28.01.2013).
5. **Zhang J.** Powerful goodness-of-fit tests based on the likelihood ratio // Journ. Royal Stat. Soc. Ser. B. 2002. **64**, Is. 2. P. 281–294.
6. **Zhang J., Wub Yu.** Likelihood-ratio tests for normality // Computat. Stat. & Data Anal. 2005. **49**, N 3. P. 709–721.
7. **Zhang J.** Powerful two-sample tests based on the likelihood ratio // Technometrics. 2006. **48**, N 1. P. 95–103.
8. **Arsham H.** Kuiper's P-value as a measuring tool and decision procedure for the goodness-of-fit test // Journ. Appl. Stat. 1988. **15**, Is, 2. P. 131–135.
9. **Stephens M. A.** The goodness-of-fit statistic V_N: distribution and significance points // Biometrika. 1965. **52**, N 3–4. P. 309–321.
10. **Stephens M. A.** EDF statistics for goodness of fit and some comparisons // Journ. Amer. Stat. Assoc. 1974. **69**, N 347. P. 730–737.
11. **Большев Л. Н., Смирнов Н. В.** Таблицы математической статистики. М.: Наука, 1983. 416 с.
12. **Pearson E. S., Hartley H. O.** Biometrika Tables for Statisticians. Cambridge: Cambridge University Press, 1972. Vol. 2. 385 p.
13. **Stephens M. A.** Use of the Kolmogorov — Smirnov, Cramer — von Mises and related statistics without extensive tables // Journ. Royal Stat. Soc. Ser. B. 1970. **32**, N 1. P. 115–122.
14. **Koziol J. A.** A modification of Watson's statistic for goodness-of-fit // Commun. in Stat. — Theory and Methods. 1989. **18**, N 10. P. 3739–3747.
15. **Nikitin Ya. Yu.** Bahadur effectiveness of Watson — Darling goodness-of-fit tests // Journ. Math. Sci. 1988. **43**, N 6. P. 2833–2838.
16. **Козлова А. В., Лемешко Б. Ю.** Исследование распределений статистик и мощности непараметрических критериев согласия, предложенных Jin Zhang // Матер. Рос. НТК «Информатика и проблемы телекоммуникаций». Новосибирск, 2007. Т. 1. С. 136–139.
17. **Лемешко Б. Ю., Лемешко С. Б., Постовалов С. Н., Чимитова Е. В.** Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход. Новосибирск: Изд-во НГТУ, 2011. 888 с.
18. **Лемешко Б. Ю., Лемешко С. Б., Постовалов С. Н.** Мощность критериев согласия при близких альтернативах // Измер. техника. 2007. № 2. С. 22–27.
19. **Лемешко Б. Ю., Лемешко С. Б., Постовалов С. Н.** Сравнительный анализ мощности критериев согласия при близких конкурирующих гипотезах. I. Проверка простых гипотез // Сиб. журн. индустр. мат. 2008. **11**, № 2(34). С. 96–111.
20. **Лемешко Б. Ю., Лемешко С. Б., Постовалов С. Н.** Сравнительный анализ мощности критериев согласия при близких альтернативах. II. Проверка сложных гипотез // Сиб. журн. индустр. мат. 2008. **11**, № 4(36). С. 78–93.
21. **Лемешко Б. Ю., Чимитова Е. В.** О выборе числа интервалов в критериях согласия типа χ^2 // Заводская лаборатория. Диагностика материалов. 2003. **69**, № 1. С. 61–67.
22. **Никулин М. С.** О критерии хи-квадрат для непрерывных распределений // Теория вероятностей и ее применение. 1973. **18**, № 3. С. 675–676.
23. **Никулин М. С.** Критерий хи-квадрат для непрерывных распределений с параметрами сдвига и масштаба // Теория вероятностей и ее применение. 1973. **18**, № 3. С. 583–591.

24. **Rao K. C., Robson D. S.** A chi-squared statistic for goodness-of-fit tests within the exponential family // Commun. Stat. 1974. **3**, N 12. P. 1139–1153.

25. **Kac M., Kiefer J., Wolfowitz J.** On tests of normality and other tests of goodness of fit based on distance methods // Ann. Math. Stat. 1955. **26**, N 1. P. 189–211.

26. **Darling D. A.** The Cramer — Smirnov test in the parametric case // Ann. Math. Stat. 1955. **26**, N 1. P. 1–20.

27. **Lilliefors H. W.** On the Kolmogorov — Smirnov test for normality with mean and variance unknown // Journ. Amer. Stat. Assoc. 1967. **62**, N 318. P. 399–402.

28. **Durbin J.** Weak convergence of the sample distribution function when parameters are estimated // Ann. Stat. 1973. **1**, N 2. P. 279–290.

29. **Мартынов Г. В.** Критерии омега-квадрат. М.: Наука, 1978. 80 с.

30. **Тюрин Ю. Н.** О предельном распределении статистик Колмогорова — Смирнова для сложной гипотезы // Изв. АН СССР. Сер. Математическая. 1984. **48**, № 6. С. 1314–1343.

31. **Лемешко Б. Ю., Постовалов С. Н.** О распределениях статистик непараметрических критериев согласия при оценивании по выборкам параметров наблюдаемых законов // Заводская лаборатория. 1998. **64**, № 3. С. 61–72.

32. **Лемешко Б. Ю., Постовалов С. Н.** Применение непараметрических критериев согласия при проверке сложных гипотез // Автометрия. 2001. № 2. С. 88–102.

33. **Лемешко Б. Ю., Постовалов С. Н.** Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Методические рекомендации. Ч. II. Непараметрические критерии. Новосибирск: Изд-во НГТУ, 1999. 85 с.

34. **Р 50.1.037-2002.** Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Ч. II. Непараметрические критерии. М.: Изд-во стандартов, 2002. 64 с.

35. **Лемешко Б. Ю., Постовалов С. Н., Французов А. В.** К применению непараметрических критериев согласия для проверки адекватности непараметрических моделей // Автометрия. 2002. № 2. С. 3–14.

36. **Лемешко Б. Ю., Маклаков А. А.** Непараметрические критерии при проверке сложных гипотез о согласии с распределениями экспоненциального семейства // Автометрия. 2004. **40**, № 3. С. 3–20.

37. **Лемешко С. Б., Лемешко Б. Ю.** Распределения статистик непараметрических критериев согласия при проверке гипотез относительно бета-распределений // ДАН ВШ России. 2007. № 2(9). С. 6–16.

38. **Лемешко Б. Ю., Лемешко С. Б.** Модели распределений статистик непараметрических критериев согласия при проверке сложных гипотез с использованием оценок максимального правдоподобия. Ч. I // Измер. техника. 2009. № 6. С. 3–11.

39. **Лемешко Б. Ю., Лемешко С. Б.** Модели распределений статистик непараметрических критериев согласия при проверке сложных гипотез с использованием оценок максимального правдоподобия. Ч. II // Измер. техника. 2009. № 8. С. 17–26.

40. **Лемешко Б. Ю., Лемешко С. Б., Никулин М. С., Саандиа Н.** Моделирование распределений статистик непараметрических критериев согласия при проверке сложных гипотез относительно обратного гауссовского закона // АиТ. 2010. № 7. С. 83–102.

41. **Lemeshko B. Yu., Lemeshko S. B., Postovalov S. N.** Statistic distribution models for some nonparametric goodness-of-fit tests in testing composite hypotheses // Commun. in Stat. — Theory and Methods. 2010. **39**, N 3. P. 460–471.

42. **Lemeshko B. Yu., Lemeshko S. B., Akushkina K. A. et al.** Inverse Gaussian model and its applications in reliability and survival analysis // Mathematical and Statistical Models and

Methods in Reliability. Applications to Medicine, Finance, and Quality Control /Eds. V. Rykov, N. Balakrishnan, M. Nikulin. Ser. "Statistics for Industry and Technology". Boston: Birkhäuser, 2011. P. 433–453.

43. **Lemeshko B. Yu., Lemeshko S. B.** Models of statistic distributions of nonparametric goodness-of-fit tests in composite hypotheses testing for double exponential law cases // Commun. in Stat. — Theory and Methods. 2011. **40**, N 16. P. 2879–2892.

44. **Lemeshko B. Yu., Lemeshko S. B.** Construction of statistic distribution models for nonparametric goodness-of-fit tests in testing composite hypotheses: The computer approach // Quality Technol. & Quantitative Management. 2011. **8**, N 4. P. 359–373.

45. **Lemeshko B. Yu., Lemeshko S. B., Rogozhnikov A. P.** Real-time studying of statistic distributions of non-parametric goodness-of-fit tests when testing complex hypotheses // Proc. of the International Workshop "Applied Methods of Statistical Analysis. Simulations and Statistical Inference". Novosibirsk, Russia, 20–22 Sept., 2011. P. 19–27.

46. **ISW** — Программная система статистического анализа одномерных случайных величин. URL: http://www.ami.nstu.ru/~headrd/ISW.htm (дата обращения 25.12.2013).

47. **Лемешко Б. Ю., Чимитова Е. В., Плешкова Т. А.** Проверка простых и сложных гипотез о согласии по цензурированным выборкам // Науч. вестн. НГТУ. 2010. № 4(41). С. 13–28.

48. **Лемешко Б. Ю., Чимитова Е. В., Ведерникова М. А.** Модифицированные критерии согласия Колмогорова, Крамера — Мизеса — Смирнова и Андерсона — Дарлинга для случайно цензурированных выборок. Ч. 1 // Науч. вестн. НГТУ. 2012. № 4(49). С. 12–19.

49. **Галанова Н. С., Лемешко Б. Ю., Чимитова Е. В.** Применение непараметрических критериев согласия к проверке адекватности моделей ускоренных испытаний // Автометрия. 2012. **48**, № 6. С. 53–68.

Поступила в редакцию 31 января 2013 г.