519.24

Сравнительный анализ критериев проверки гипотезы о равномерности закона

Б. Ю. ЛЕМЕШКО, П. Ю. БЛИНОВ

Новосибирский государственный технический университет, Новосибирск, Россия, e-mail: lemeshko@ami.nstu.ru

Рассмотрено множество специальных критериев, предназначенных для проверки гипотез о принадлежности наблюдений равномерному закону. Исследованы распределения статистик критериев и их мощность относительно различных конкурирующих гипотез. Выявлены достоинства и недостатки отдельных критериев. Показано, что значительная часть критериев, традиционно используемых при проверке гипотез о равномерности, оказывается смещённой относительно некоторого вида конкурирующих гипотез. Подчеркивается, что специальные критерии проверки равномерности не имеют явных преимуществ перед непараметрическими критериями согласия, применяемыми для проверки равномерности.

Ключевые слова: *равномерный закон, проверка гипотез, статистический критерий, мощность критерия.*

The set of special tests intended for testing uniformity are considered. Distributions of test statistics, power of tests under different competing hypotheses are studied. Considered test are ranked by test power. Advantages and disadvantages of individual tests were shown. It has been shown that large parts of the tests traditionally used for testing uniformity have the bias under some kind of competing hypotheses. Underlines that special uniformity tests haven't clear advantage over nonparametric goodness-of-fit tests used for testing uniformity in general.

Key words: *uniform distribution, hypothesis testing, test statistic, test power.*

В прикладной математической статистике равномерный закон распределения вероятностей занимает важное место. Иногда его используют в качестве модели для описания ошибок измерений некоторых приборов или измерительных систем.

Функция распределения вероятностей равномерного закона на интервале [0, 1] имеет вид $F(x) = x$. Гипотезу о принадлежности выборки $X_1, X_2, \dots, X_n$ независимых наблюдений случайной величины X равномерному закону можно записать в виде H_0: $F(x) = x$, $x \in [0, 1]$. В большинстве критериев проверки гипотезы о равномерности на указанном интервале опираются на оценки порядковых статистик величины X (на элементы $x_{(i)}$ вариационного ряда $0 < x_{(1)} < x_{(2)} < \dots < x_{(n)} < 1$, построенного по $X_1, X_2, \dots, X_n$). В дальнейшем в выражениях статистик критериев будем использовать обозначения $U_{(i)} = x_{(i)}$, $i = \overline{1, n}$, $U_0 = 0$, $U_{n+1} = 1$.

Как правило, все критерии ориентированы на проверку *простой* гипотезы H_0 на интервале [0, 1]. В случае необходимости проверки гипотезы о принадлежности выборки $X_1, X_2, \dots, X_n$ равномерному закону на интервале $[a, b]$ (с параметрами сдвига a и масштаба $b-a$) для использования всех критериев равномерности элементы $x_{(i)}$ вариационного ряда $a < x_{(1)} < x_{(2)} < \dots < x_{(n)} < b$, построенного по выборке $X_1, X_2, \dots, X_n$, преобразуют в соответствующие (требуемые в критериях) порядковые статистики следующим образом: $U_{(i)} = (x_{(i)} - a)/(b - a)$, $i = \overline{1, n}$, $U_0 = 0$, $U_{n+1} = 1$. Весь дальнейший порядок применения критериев проверки равномерности остаётся неизменным (как и на интервале [0, 1]).

При проверке *сложной* гипотезы о равномерности вида H_0: $F(x) = (x - a)/(b - a)$, $x \in [a, b]$ где a, b неизвестны и должны быть найдены по этой же выборке, поступают следующим образом. По элементам вариационного ряда $x_{(1)} < x_{(2)} < \dots < x_{(n)}$, построенного по выборке $X_1, X_2, \dots, X_n$, находят оценки параметров

$$\hat{a} = x_{(1)} - \left(x_{(n)} - x_{(1)}\right)/(n-1); \; \hat{b} = x_{(n)} + \left(x_{(n)} - x_{(1)}\right)/(n-1).$$

Очевидно, что проверка сложной гипотезы о принадлежности $X_1, X_2, \dots, X_n$ равномерному закону на интервале $[\hat{a}, \hat{b}]$, полученному по данной выборке, эквивалентна проверке простой гипотезы о принадлежности части выборки меньшего объёма $n-2$ (соответствующей ряду $x_{(2)} < x_{(3)} < \dots < x_{(n-1)}$) равномерному закону на интервале $[x_{(1)}, x_{(n)}]$, который соответствует размаху выборки. В этом случае при использовании всех рассматриваемых критериев требуемые значения порядковых статистик находятся в соответствии с соотношениями

$$U_{i-1} = \left(x_{(i)} - x_{(1)}\right)/\left(x_{(n)} - x_{(1)}\right);$$

$$i = \overline{2, (n-1)}; \; U_0 = 0; \; U_{n-1} = 1.$$

Для проверки гипотезы о принадлежности выборки равномерному закону предложено множество статистических критериев, которые можно разделить на два подножества. В первое входят специальные критерии, ориентированные на проверку равномерности, а во второе — различные критерии согласия, которые также можно применять для проверки равномерности. Существование множества критериев ставит практиков перед сложным выбором, так как имеющаяся в публикациях информация не позволяет однозначно отдать предпочтение какому-то определенному критерию.

Настоящая работа посвящена сравнительному анализу специальных критериев, в подмножестве которых можно выделить три основные группы критриев проверки равномерности.

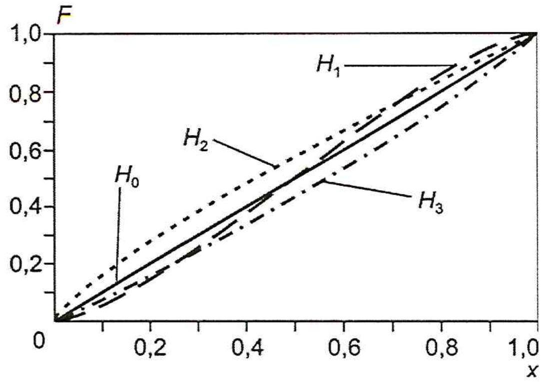

Рис. 1. Функции распределения вероятностей, соответствующие конкурирующим гипотезам

Статистики критериев первой группы предусматривают использование разностей последовательных значений вариационного ряда

$$D_i = U_i - U_{i-1},$$

где $i=\overline{1,(n+1)}$, n — объем выборки; $U_0 = 0$; $U_{n+1} = 1$.

Ко второй группе относят различные модификации критериев, использующие разности оценок порядковых статистик, соответствующих анализируемой выборке, и, например, математических ожиданий этих порядковых статистик.

Третью группу составляют, так называемые, энтропийные критерии, опирающиеся на различные оценки энтропии.

Рассматриваемые конкурирующие гипотезы. С результатами проверки гипотез связывают ошибки двух видов: при ошибке первого рода гипотезу H_0 отклоняют, когда она верна; а ошибка второго рода состоит в том, что принимают (не отклоняют) гипотезу H_0, в то время как справедлива конкурирующая гипотеза H_1. Уровень значимости α задает вероятность ошибки первого рода.

Обычно, используя критерии проверки гипотез, не рассматривают конкретную конкурирующую гипотезу. В таком случае при проверке гипотез о виде закона можно считать, что конкурирующая гипотеза имеет вид H_1: $F(x)\neq F(x, 0)$. Если же гипотеза H_1 задана и имеет вид H_1: $F(x)=F_1(x, \theta_0)$ то задание величины α для используемого критерия проверки гипотез определяет и вероятность ошибки второго рода β.

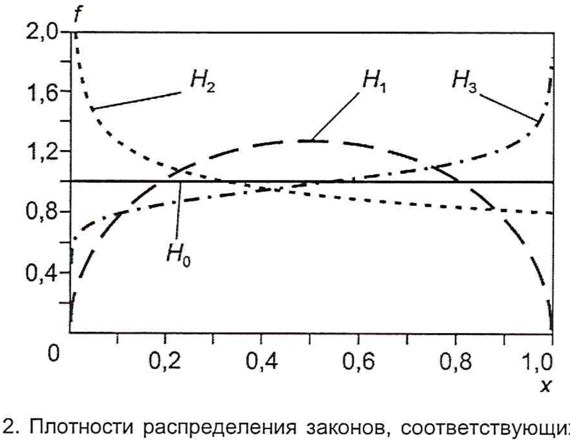

Рис. 2. Плотности распределения законов, соответствующих конкурирующим гипотезам

Ошибка второго рода заключается в том, что не отклоняется гипотеза H_0, когда на самом деле справедлива гипотеза H_1. Мощность критерия является разностью $1-\beta$. Очевидно, что чем выше мощность используемого критерия при заданном значении α, тем лучше он различает гипотезы H_0, H_1.

Наибольший интерес представляет способность критериев различать близкие конкурирующие гипотезы. При анализе близких альтернатив удаётся выяснить тонкие моменты, характеризующие свойства критериев, выявить их принципиальные недостатки или достоинства.

В данной работе мощность всех рассмотренных критериев исследовали относительно трёх конкурирующих гипотез, которые соответствуют принадлежности наблюдаемой случайной величины семейству бета-распределений 1-го рода с функцией плотности

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

где $\mathrm{B}(\theta_0, \theta_1) = \Gamma(\theta_0)\,\Gamma(\theta_1)/\Gamma(\theta_0 + \theta_1)$ — бета-функция; $\theta_0, \theta_1 \in (0, \infty)$ — параметры формы; $\theta_2 \in (0, \infty)$ — масштабный параметр; $\theta_3 \in (-\infty, \infty)$ — параметр сдвига; $x \in [0, \theta_2]$.

Обозначим функцию бета-распределения 1-го рода при конкретных значениях параметров как $\mathrm{B_I}(\theta_0; \theta_1; \theta_2; \theta_3)$. Тогда три рассматриваемые и достаточно близкие к H_0 конкурирующие гипотезы H_1, H_2, H_3 принимают следующий вид:

$$H_1\text{: } F(x) = \mathrm{B_I}\,(1{,}5;\ 1{,}5;\ 1;\ 0),\ x \in [0,\ 1];$$

$$H_2\text{: } F(x) = \mathrm{B_I}\,(0{,}8;\ 1;\ 1;\ 0),\ x \in [0,\ 1];$$

$$H_3\text{: } F(x) = \mathrm{B_I}\,(1{,}1;\ 0{,}9;\ 1;\ 0),\ x \in [0,\ 1].$$

Функции распределения вероятностей, соответствующие рассматриваемым гипотезам, представлены на рис. 1, а плотности распределений — на рис. 2. Следует обратить внимание, что конкурирующей гипотезе H_1 соответствует закон, функция распределения которого пересекается с функцией распределения равномерного закона, а при H_2, H_3 функции распределения законов лежат выше и ниже функции равномерного. Способности различать гипотезы H_0, H_1 и гипотезы H_0, H_2 или H_3 у критериев различны.

Анализ мощности критериев относительно H_1 позволил выявить неспособность отдельных критериев при малых объёмах выборок n и малых уровнях значимости α отличать эту гипотезу от H_0, т. е. показал смещённость соответствующих критериев (мощность $1-\beta$ оказывается меньше α). Указанный недостаток оказался свойственным не только значительной части специальных критериев проверки равномерности, но и большей части непараметрических критериев согласия [1].

Результаты исследований. Как было сказано выше, множество специальных критериев, статистики которых приведены в табл. 1, можно разбить на три группы близких по свойствам критериев. К первой группе критериев, использующих разности элементов вариационного ряда, относятся критерии Шермана [2, 3], Кимбелла [4], Морана 1 [5], Морана 2 [6], критерии Кресси с уточнёнными по сравнению с [7] выражениями статистик $S_n^{(m)}$, $L_n^{(m)}$ [8], Пардо [8], Шварца [10].

Таблица 1

Критерии проверки равномерности

Критерий	Статистика
Шермана	$\omega_n = \frac{1}{2}\sum_{i=1}^{n+1}\left\|U_i - U_{i-1} - \frac{1}{n+1}\right\|$
Кимбелла	$A = \sum_{i=1}^{n+1}\left(U_i - U_{i-1} - \frac{1}{n+1}\right)^2$
Морана 1	$B = \sum_{i=1}^{n+1}\left(U_i - U_{i-1}\right)^2$
Морана 2	$M_n = -\sum_{i=1}^{n+1}\ln\left[(n+1)\left(U_i - U_{i-1}\right)\right]$
Ченга—Спиринга	$W_p = \left[\left(U_n - U_1\right)\frac{n+1}{n-1}\right]^2 \Big/ \sum_{i=1}^{n}\left(U_i - \overline{U}\right)^2$
Янга	$M = \frac{1}{l}\sum_{i=1}^{n}\min\left(D_i, D_{i+1}\right)$, $D_1 = U_1$, $D_i = U_i - U_{i-1}$, $D_{n+1} = 1 - U_n$
Гринвуда	$G = (n+1)\sum_{i=1}^{n+1}\left(U_i - U_{i-1}\right)^2$
Гринвуда—Кэсенберри—Миллера	$Q = \sum_{i=1}^{n+1}\left(U_i - U_{i-1}\right)^2 + \sum_{i=1}^{n}\left(U_{i+1} - U_i\right)\left(U_i - U_{i-1}\right)$
Шварца	$A_n^* = \frac{n}{2}\sum_{i=1}^{n}\left(\frac{U_{i+1} - U_{i-1}}{2} - \frac{1}{n}\right)^2$, $U_0 = -U_1$, $U_{n+1} = 2 - U_n$
Хегази—Грина T_1	$T_1 = \frac{1}{n}\sum_{i=1}^{n}\left\|U_i - \frac{i}{n+1}\right\|$
Хегази—Грина T_1^*	$T_1^* = \frac{1}{n}\sum_{i=1}^{n}\left\|U_i - \frac{i-1}{n-1}\right\|$
Хегази—Грина T_2	$T_2 = \frac{1}{n}\sum_{i=1}^{n}\left(U_i - \frac{i}{n+1}\right)^2$
Хегази—Грина T_2^*	$T_2^* = \frac{1}{n}\sum_{i=1}^{n}\left(U_i - \frac{i-1}{n-1}\right)^2$
Фросини	$B_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\|U_i - \frac{i-0{,}5}{n}\right\|$
Неймана—Бартона, N_2	$N_2 = \sum_{j=1}^{2}V_j^2$, $V_j = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\pi_j\left(U_i - 0{,}5\right)$, $\pi_1(y) = 2\sqrt{3}\,y$; $\pi_2(y) = \sqrt{5}\left(6y^2 - 0{,}5\right)$
Неймана—Бартона, N_3	$N_3 = \sum_{j=1}^{3}V_j^2$, $V_j = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\pi_j\left(U_i - 0{,}5\right)$, $\pi_3(y) = \sqrt{7}\left(20y^3 - 3y\right)$

Окончание таблицы 1

Критерий	Статистика
Неймана—Бартона, N_4	$N_4=\sum_{j=1}^{4}V_j^2,\ V_j=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\pi_j\left(U_i-0{,}5\right),$ $\pi_4(y)=3\left(70y^4-15y^2+0{,}375\right)$
Дудевича-ван дер Мюлена	$H(m,n)=-\frac{1}{n}\sum_{i=1}^{n}\ln\left\{\frac{n}{2m}\left(U_{i+m}-U_{i-m}\right)\right\}$, m — целое, $m\le\frac{n}{2}$; если $i+m\ge n$, то $U_{i+m}=U_n$, и если $i-m\le 1$, то $U_{i-m}=U_1$
Модификация 1 энтропийного критерия	$HY_1=-\frac{1}{n}\sum_{i=1}^{n}\ln\left(\frac{U_{i+m}-U_{i-m}}{\hat F\left(U_{i+m}\right)-\hat F\left(U_{i-m}\right)}\right),$ $\hat F\left(U_i\right)=\frac{n-1}{n(n+1)}\left(i+\frac{1}{n-1}+\frac{U_i-U_{i-1}}{U_{i+1}-U_{i-1}}\right),\ i=2,\ldots,n-1,$ $\hat F\left(U_1\right)=1-\hat F\left(U_n\right)=1/(n+1)$
Модификация 2 энтропийного критерия	$HY_2=-\sum_{i=1}^{n}\ln\left(\frac{U_{i+m}-U_{i-m}}{\hat F\left(U_{i+m}\right)-\hat F\left(U_{i-m}\right)}\right)\left(\frac{\hat F\left(U_{i+m}\right)-\hat F\left(U_{i-m}\right)}{\sum_{j=1}^{n}\left(\hat F\left(U_{j+m}\right)-\hat F\left(U_{j-m}\right)\right)}\right)$
Пардо	$E_{m,n}=-\frac{1}{n}\sum_{i=1}^{n}\frac{2m}{n\left(U_{i+m}-U_{i-m}\right)}$
Кресс_и 1	$S_n^{(m)}=\sum_{i=0}^{n+1-m}\left(U_{i+m}-U_i-\frac{m}{n+1}\right)^2$
Кресси 2	$L_n^{(m)}=-\sum_{i=0}^{n+1-m}\ln\left[\frac{n+1}{m}\left(U_{i+m}-U_i\right)\right]$

Ко второй группе критериев, где рассматриваются отклонения порядковых статистик от их математических ожиданий (от медиан и т. п.), относятся критерии Хегази—Грина со статистиками T_1, T_2 [11], Фросини [12], Янга [13], Ченга—Спиринга [14], Гринвуда [15], Гринвуда—Кэсенберри—Миллера [16], критерии Неймана—Бартона со статистиками N_2, N_3 и N_4 [17].

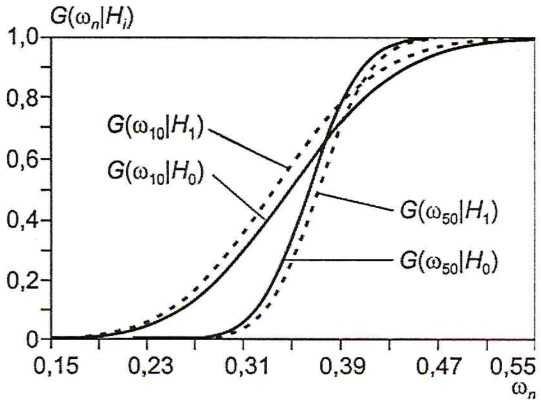

Рис. 3. Распределения $G(\omega_n|H_i)$ статистики критерия Шермана

К третьей группе относится энтропийный критерий Дудевича-ван дер Мюлена [18] и две модификации, в статистиках которых используются другие оценки энтропии [19].

При выполнении данной работы, которая явилась развитием [20, 21], методами статистического моделирования [22] были исследованы распределения статистик всех перечисленных критериев, расширены таблицы процентных точек, проверено, насколько распределения нормализованных статистик описываются соответстветствующими асимптотическими законами. Исследована мощность критериев относительно различных конкурирующих гипотез, в частности относительно H_1, H_2, H_3. Оказалось, что целый ряд из рассмотренных критериев является смещённым относительно конкурирующей гипотезы H_1.

Для доказательства факта смещённости на рис. 3 показаны распределения $G(\omega_n|H_i)$ статистики критерия Шермана, соответствующие справедливости гипотез H_0, H_1 при объёмах выборок n = 10, n = 50.

Правосторонний критерий и проверяемая гипотеза отклоняются при повышенных значениях статистики. Распре-

Таблица 2

Упорядоченность критериев равномерности по мощности относительно конкурирующих гипотез H_1, H_2, H_3

Относительно H_1	$1-\beta$			Относительно H_2	$1-\beta$			Относительно H_3	$1-\beta$		
	$n = 10$	$n = 50$	$n = 100$		$n = 10$	$n = 50$	$n = 100$		$n = 10$	$n = 50$	$n = 100$
Модификация 2 энтропийного критерия	0,265	0,704	0,883	Хегази—Грина T_1	0,178	0,397	0,610	Хегази—Грина T_1	0,149	0,330	0,522
Неймана—Бартона N_2	0,116	0,525	0,837	Фросини	0,171	0,389	0,603	Фросини	0,148	0,330	0,522
Крессии 2	0,201	0,644	0,820	Хегази—Грина T_2	0,175	0,389	0,602	Хегази—Грина T_1^*	0,145	0,327	0,520
Дудевича-ван дер Мюлена	0,254	0,601	0,790	Неймана—Бартона N_2	0,177	0,333	0,597	Хегази—Грина T_2	0,147	0,322	0,508
Модификация 1 энтропийного критерия	0,254	0,600	0,789	Хегази—Грина T_1^*	0,160	0,378	0,595	Хегази—Грина T_2^*	0,144	0,319	0,506
Неймана—Бартона N_3	0,087	0,431	0,766	Хегази—Грина T_2^*	0,159	0,371	0,585	Неймана—Бартона N_2	0,137	0,279	0,447
Неймана—Бартона N_4	0,072	0,371	0,739	Неймана—Бартона N_3	0,177	0,370	0,577	Неймана—Бартона N_3	0,133	0,258	0,416
Ченга—Спиринга	0,109	0,388	0,722	Неймана—Бартона N_4	0,177	0,359	0,557	Неймана—Бартона N_4	0,130	0,240	0,381
Шварца	0,154	0,398	0,583	Пардо	0,121	0,296	0,463	Пардо	0,118	0,208	0,291
Хегази—Грина T_1^*	0,101	0,226	0,443	Модификация 1 энтропийного критерия	0,097	0,189	0,328	Дудевича-ван дер Мюлена	0,115	0,191	0,275
Хегази—Грина T_2^*	0,097	0,210	0,409	Дудевича-ван дер Мюлена	0,097	0,188	0,327	Модификация 1 энтропийного критерия	0,115	0,191	0,275
Пардо	0,168	0,255	0,408	Крессии 1	0,157	0,239	0,314	Модификация 2 энтропийного критерия	0,114	0,185	0,267
Фросини	0,072	0,177	0,384	Модификация 2 энтропийного критерия	0,095	0,153	0,266	Крессии 2	0,123	0,170	0,226
Хегази—Грина T_1	0,054	0,133	0,322	Гринвуда—Кэсенберри—Миллера	0,159	0,204	0,244	Крессии 1	0,118	0,167	0,218
Хегази—Грина T_2	0,060	0,137	0,308	Шварца	0,136	0,184	0,226	Шварца	0,129	0,172	0,206
Гринвуда—Кэсенберри—Миллера	0,036	0,142	0,290	Крессии 2	0,120	0,140	0,217	Гринвуда—Кэсенберри—Миллера	0,119	0,155	0,186
Кимбелла	0,059	0,160	0,279	Шермана	0,147	0,177	0,204	Кимбелла	0,116	0,142	0,165
Морана 1	0,059	0,160	0,279	Кимбелла	0,143	0,174	0,201	Морана 1	0,116	0,142	0,165
Гринвуда	0,059	0,160	0,279	Морана 1	0,143	0,174	0,201	Гринвуда	0,116	0,142	0,165
Шермана	0,065	0,147	0,215	Гринвуда	0,143	0,174	0,201	Шермана	0,117	0,137	0,154

Окончание таблицы 2

Относительно H_1	$1-\beta$			Относительно H_2	$1-\beta$			Относительно H_{13}	$1-\beta$		
	$n = 10$	$n = 50$	$n = 100$		$n = 10$	$n = 50$	$n = 100$		$n = 10$	$n = 50$	$n = 100$
Кресси 1	0,037	0,082	0,187	Морана 2	0,150	0,172	0,193	Морана 2	0,116	0,131	0,143
Морана 2	0,072	0,135	0,187	Ченга—Спиринга	0,106	0,134	0,168	Ченга—Спиринга	0,100	0,105	0,106
Янга	0,090	0,108	0,115	Янга	0,108	0,108	0,108	Янга	0,102	0,104	0,104

деление статистики $G(\omega_{10}|H_1)$, имеющее место при справедливости H_1, сдвинуто относительно $G(\omega_{10}|H_0)$ не вправо, а влево. Следовательно, мощность $1-\beta$ оказывается меньше соответствующего α. С ростом n смещённость исчезает (см. $G(\omega_{50}|H_0)$ и $G(\omega_{50}|H_1)$ на рис. 3).

Выражения для статистик рассмотренных критериев проверки равномерности приведены в табл. 1. В табл. 2 специальные критерии проверки равномерности упорядочены по убыванию мощности относительно конкурирующих гипотез H_1, H_2, H_3 (по величине мощности $1-\beta$, проявленной при n=100 и уровне значимости $\alpha = 0,1$). В табл. 2 также приведены оценки мощности критериев при n=10 и n=50.

Некоторый сравнительный анализ критериев равномерности проводился в [23]. В [19] энтропийный критерий сравнивали с непараметрическими критериями согласия. Разрозненные оценки мощности можно найти и в других работах.

Типичным недостатком большей части применяемых критериев равномерности, представленных в табл. 1, является отсутствие информации о распределениях статистик и вынужденная небходимость использования таблиц критических значений (процентных точек), что в принципе можно обойти при существовании соответствующего программного обеспечения [8]. Другой недостаток большей части критериев — смещённость относительно конкурирующих гипотез вида H_1 (при малых объёмах выборок n и малых уровнях значимости α). В табл. 2 тёмным цветом выделены критерии, которые при малых n обладают ярко выраженной смещённостью относительно гипотезы H_1. В меньшей степени смещённость относительно проявляется у критериев Неймана—Бартона со статистиками H_2, H_3. Этот недостаток не отмечен только для некоторых критериев: для энтропийного критерия Дудевича-ван дер Мюлена и его модификаций, для критериев Ченга—Спиринга, Шварца и Пардо.

Все модификации критериев, использующие в качестве статистик различные оценки энтропии [18, 19], показали высокую мощность относительно конкурирующей гипотезы H_1. В то же время относительно гипотез вида H_2, H_3 оценки мощности этих критериев более скромные. Надо отметить, что только у этих критериев при малых n наблюдаются признаки смещённости относительно гипотезы H_2.

Критерий Неймана—Бартона со статистикой N_2 показывает высокую мощность относительно H_1 и сравнительно высокие результаты относительно H_2, H_3. Стабильно неплохую способность отличать конкурирующие гипотезы от равномерного закона демонстрируют критерии Хегази—Грина и Фросини. Низкую мощность демонстрируют критерии, в статистиках которых суммируются модули или квадраты разностей $U_i - U_{i-1}$ значений последовательных порядковых статистик (критерии Шермана, Кимбелла, Морана, Гринвуда, Гринвуда—Кэсенберри—Миллера). Особенно низкую мощность относительно всех трёх рассматриваемых в работе гипотез показал критерий Янга [13], что свидетельствует о крайней неудачности попытки использования соответствующей статистики в критерии проверки гипотезы о равномерности. Можно предположить, что столь же неудачной идеей окажется применение такой статистики в любом критерии, предназначенном для проверки согласия наблюдаемой выборки с некоторым конкретным законом распределения.

На основании исследования свойств множества критериев, используемых при проверке равномерности, подготовлено руководство по применению [8]. Когда для проверки гипотезы о принадлежности анализируемой выборки некоторому конкретному закону распределения разработано множество специальных критериев, то среди этого множества, как правило, найдутся критерии, предпочтительность применения которых при ограниченных объёмах выборок связано с заметным преимуществом в мощности, например, по сравнению с общими критериями согласия. В данном случае (при проверке равномерности) такого преимущества относительно непараметрических критериев согласия не наблюдается: стабильно показывают себя критерии Жанга со статистиками Z_A, Z_C и критерий Андерсона—Дарлинга [8].

Из анализа свойств всего множества критериев, которые можно использовать для проверки гипотезы о принадлежности выборки равномерному закону, вытекает, что корректного использования какого-то одного из критериев для формирования «надежного» статистического вывода часто может оказаться недостаточно. Для бóльшей объективности статистических выводов предпочтительней воспользоваться некоторым рядом критериев, обладающих определёнными достоинствами. Использование совокупности критериев, опирающихся на различные меры отклонения эмпирического распределения от теоретического, повышает качество статистических выводов.

Смещённость относительно некоторых близких конкурирующих гипотез при малых проявляют не только критерии равномерности. Аналогичным недостатком обладают некоторые критерии, используемые для проверки нормальности [24—26].

Работа выполнена при поддержке Минобрнауки РФ в рамках государственной работы «Обеспечение проведения научных исследований» и проектной части государственного задания (проект № 2.541.2014К).

Литература

1. **Лемешко Б. Ю., Блинов П. Ю., Лемешко С. Б.** Смещённость непараметрических критериев согласия относительно некоторых пар конкурирующих гипотез // Измерительная техника. 2016. № 5. С. 16—20.

2. **Sherman B.** A random variable related to the spacing of sample values / B. Sherman // The Annals of Mathematical Statistics. 1950. V. 21. N. 3. P. 339—361.

3. **Sherman B.** Percentiles of the w_n statistic / B. Sherman // The Annals of Mathematical Statistics. 1957. V. 2. N. 1. P. 25—261.

4. **Kimball B. F.** Some basic theorems for developing tests of fit for the case of the non-parametric probability distribution function / B. F. Kimball // The Annals of Mathematical Statistics. 1947. V. 18. N. 1. P. 540—548.

5. **Moran P. A. P.** The random division of an intervals / P. A. P. Moran // J. R. Statist. Soc. 1947. Ser. B. V. 9. N. 1. P. 92—98.

6. **Moran P. A. P.** The random division of an intervals. II / P. A. P. Moran // J. R. Statist. Soc. 1951. Ser. B. V. 13. N. 2. P. 147—150.

7. **Cressie N.** An optimal statistic based on higher order gaps // Biometrika. 1979. V. 66. P. 619—627.

8. **Лемешко Б. Ю.** Критерии проверки отклонения распределения от равномерного закона: Руководство по применению / Б. Ю. Лемешко, П. Ю. Блинов. М.: ИНФРА-М, 2015.

9. **Pardo M. C.** A test for uniformity based on informational energy // Statist. Papers. 2003. V. 44. P. 521—534.

10. **Swartz T.** Goodness-of-fit tests using Kullback-Leibler information // Communications in Statistics — Theory and Methods. 1992. V. 21. P. 711—729.

11. **Hegazy Y. A. S.** Some new goodness-of-fit tests using order statistics / Y. A. S. Hegazy, J. R. Green // Appl. Statist. 1975. V. 24. N. 3. P. 299—308.

12. **Frosini B. V.** On the distribution and power of goodness-of-fit statistic with parametric and nonparametric applications, «Goodness-of-fit» / Ed. by Revesz P., Sarkadi K., Sen P. K. // Amsterdam—Oxford—N. Y.: North-Holland Publ. Comp. 1987. P. 133—154.

13. **Young D. L.** The linear nearest neighbour statistic / D. L. Young // Biometrika. 1982. V. 69. N. 2. P. 477—480.

14. **Cheng S. W.** A test to Identify the uniform distribution with applications / S. W. Cheng, F. A. Spiring // IEEE Trans. Reliability. 1987. V. R-36. P. 98—105.

15. **Greenwood V.** The statistical study of Infection disease / V. Greenwood // J. R. Statist. Soc. 1946. Ser. A. V. 109. P. 257—261.

16. **Quesenberry C. P.** Power studies of some tests for uniformity / C. P. Quesenberry, F. L. Miller // J. Statist. Computation and Simulation. 1977. V. 5. P. 169—191.

17. **Neyman J.** Smooth» tests for goodness-of-fit // Scandinavisk Aktuarietidskrift. 1937. V. 20. P. 149—199.

18. **Dudewics E. J., van der Meulen E. C.** Entropy-based test of uniformity // J. Amer. Statist. Assoc.1981. V. 76. N. 376. P. 967—974.

19. **Zamanzade E.** Testing uniformity based on new entropy estimators // J. Statist. Computation and Simulation. 2014.

20. **Blinov P. Yu., Lemeshko B. Yu.** A review of the properties of tests for uniformity // 12th Int. Conf. on Actual Problems of Electronics Instrument Engineering (APEIE) 34006 Proc. 2014. V. 1. P. 540—547.

21. **Блинов П. Ю., Лемешко Б. Ю.** О мощности критериев, используемых для проверки гипотез о принадлежности выборок равномерному закону // Обработка информационных сигналов и математическое моделирование: материалы Российской НТК, Новосибирск. 2013. С. 35—38.

22. **Лемешко Б. Ю.** Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. Новосибирск: Изд-во НГТУ, 2011.

23. **Marhuenda Y., Morales D., Pardo M. C.** A comparison of uniformity tests // Statistics. 2005. V. 39. N. 4. P. 315—327.

24. **Лемешко Б. Ю., Лемешко С. Б.** Сравнительный анализ критериев проверки отклонения распределения от нормального закона // Метрология. 2005. № 2. С. 3—24.

25. **Лемешко Б. Ю., Рогожников А. П.** Исследование особенностей и мощности некото—24.

26. **Лемешко Б. Ю.** Критерии проверки отклонения распределения от нормального закона: Руководство по применению М.: ИНФРА-М, 2015.

Дата принятия: 27.07.2015.