Исследование методами статистического моделирования свойств некоторых критериев нормальности

Борис Ю. Лемешко[1], Андрей П. Рогожников[1]
[1]*Новосибирский государственный технический университет*

Аннотация **– Исследованы достоинства и недостатки, оценена мощность различных критериев проверки отклонения от нормального закона (критериев Фроцини, Хегази-Грина, Шпигельхальтера, Гири и Дэвида-Хартли-Пирсона).**

Ключевые слова **– Критерии нормальности, мощность критерия, нормальный закон**

I. ВВЕДЕНИЕ

ЗНАЧИТЕЛЬНАЯ часть методов статистических анализа базируется на предположении о принадлежности исходных данных нормальному закону распределения. Поэтому использование таких методов предполагает обязательную проверку гипотезы о нормальности.

В российском стандарте ГОСТ Р ИСО 5479-2002 «Статистические методы. Проверка отклонения распределения вероятностей от нормального распределения» [1] закреплено применение различных критериев. Однако в нём не даётся рекомендаций относительно того, какой критерий наиболее предпочтителен и имеет наибольшую мощность при заданном объёме выборки или против определённых конкурирующих гипотез.

В [2] критерии, упоминаемые в [1], исследованы наряду с некоторыми другими и даны рекомендации по применению. Там же [2] были отмечены определённые недостатки некоторых критериев, в частности, критериев Шапиро–Уилка и Эппса–Палли. В литературе уделено не очень много внимания сравнительному анализу различных критериев нормальности. Можно сказать, что вопрос о выборе наилучшего критерия нормальности до сих пор остаётся открытым, так рекомендации или противоречивы, или предвзяты.

В данной работе проводится подробное исследование свойств критериев, предложенных Фроцини [3,4], Хегази и Грином [5], Шпигельхальтером [6], Гири [7] и Дэвидом, Хартли и Пирсоном [8]. Цель этих исследований заключалась в выявлении достоинств и недостатков данных критериев, в сравнительном анализе их мощности, сопоставление с мощностью критериев Шапиро-Уилка и Эппса-Палли и с мощностью критериев согласия.

II. КРИТЕРИИ И ГИПОТЕЗЫ

1. *Рассматриваемые критерии*

Рассматриваемые критерии нормальности основаны на следующих статистиках.

Статистика Фроцини:

$$B_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left| \Phi(z_i) - \frac{i-0.5}{n} \right|, \qquad (1)$$

где $x_{(i)}$ - порядковые статистики,

$$z_i = \frac{x_{(i)} - \bar{x}}{s}, \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2,$$

а $\Phi(z)$ – функция распределения стандартного нормального закона $N(0,1)$.

Статистики Хегази-Грина:

$$T_1 = \frac{1}{n}\sum_{i=1}^{n} |z_i - \eta_i|, \qquad (2)$$

$$T_2 = \frac{1}{n}\sum_{i=1}^{n} \{z_i - \eta_i\}^2, \qquad (3)$$

где z_i, $x_{(i)}$ и $\bar{x}$ определяются так же, как для статистики Фроцини,

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2,$$

η_i - математическое ожидание i-й порядковой статистики стандартного нормального закона, которое можно найти как

$$\eta_i = \Phi^{-1}\left(\frac{i}{n+1}\right).$$

Проверяемая гипотеза отклоняется при больших значениях статистики.

Статистика Гири:

$$d = \frac{1}{ns}\sum_{i=1}^{n} |x_i - \bar{x}|, \qquad (4)$$

где $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $s^2 = \frac{1}{n}\sum_{i=1}^{n}\ x_i - \bar{x}\ ^2$. Статистика распределена асимптотически нормально при $n \geq 50$ со средним

$$M\ \ d\ \ = \frac{2}{\sqrt{\pi\ \ n-1}}\Gamma\left(\frac{n+1}{2}\right)\Big/\Gamma\left(\frac{n}{2}\right)$$

и дисперсией

$$D\ \ d\ \ = \frac{1}{n}\left\{1 + \frac{2}{\pi}\left[\sqrt{n\ \ n-2}\ + \arcsin\left(\frac{1}{n-1}\right)\right]\right\} -$$

$$-\frac{n-1}{\pi}\left[\Gamma\left(\frac{n-1}{2}\right)\Big/\Gamma\left(\frac{n}{2}\right)\right]^2.$$

Критерий является двусторонним, и гипотеза о нормальности принимается, если

$$d\left(\frac{\alpha}{2}\right) \leq d \leq d\left(1 - \frac{\alpha}{2}\right),$$

где $d(a) = M(d) - u_a D(d)$, u_a – a-квантиль стандартного нормального распределения.

Статистика Дэвида–Хартли–Пирсона:

$$U = \frac{R}{s}, \tag{5}$$

где $R = x_{\max} - x_{\min}$ – размах выборки, $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ – несмещённая оценка дисперсии. Гипотеза о нормальности распределения отвергается, если $U < U_1(\alpha)$ или $U > U_2(\alpha)$ (α – уровень значимости).

Статистика Шпигельхальтера является комбинацией статистик критериев Гири и Дэвида–Хартли–Пирсона, и имеет вид:

$$T' = \ (C_n U)^{-(n-1)} + g^{-(n-1)}\ ^{\frac{1}{n-1}}, \tag{6}$$

где $C_n = \frac{1}{2n}(n!)^{\frac{1}{n-1}}$, U – статистика (5) критерия Дэвида–Хартли–Пирсона, $g = \frac{d}{\sqrt{(n-1)/n}}$, d – статистика (4) критерия Гири. Гипотеза о нормальности отклоняется при больших значениях статистики T'.

2. *Конкурирующие гипотезы*

В настоящей работе при исследовании мощности рассматриваются те же конкурирующие гипотезы, что и в [2]. Проверяемой гипотезе H_0 соответствует нормальный закон с плотностью

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

с параметром масштаба $\theta_1 = 1$ и параметром сдвига $\theta_0 = 0$. Конкурирующей гипотезе H_1 соответствует распределение семейства

$$f(x) = \frac{\theta_2}{2\theta_1\Gamma(1/\theta_2)}\exp\left\{-\left(\frac{|x-\theta_0|}{\theta_1}\right)^{\theta_2}\right\} \tag{7}$$

с параметром формы $\theta_2 = 4$, параметром масштаба $\theta_1 = 1$ и параметром сдвига $\theta_0 = 0$; конкурирующей гипотезе H_2 – распределение семейства (7) с параметром формы $\theta_2 = 1$ (распределение Лапласа), параметрами масштаба $\theta_1 = 1$ и сдвига $\theta_0 = 0$; H_3 – логистическое распределение с плотностью

$$f(x) = \frac{\pi}{\theta_1\sqrt{3}}\exp\left\{-\frac{\pi(x-\theta_0)}{\theta_1\sqrt{3}}\right\}\Bigg/$$

$$\Bigg/\left[1 + \exp\left\{-\frac{\pi(x-\theta_0)}{\theta_1\sqrt{3}}\right\}\right]^2$$

и параметрами $\theta_0 = 0$ и $\theta_1 = 1$. Рисунок, показывающий плотности распределений, соответствующих каждой из гипотез H_i, приведен в [2].

3. *Методика исследования*

При исследовании распределений, построении процентных точек статистик критериев и оценке мощности критериев относительно различных конкурирующих гипотез в данной работе использовалась методика статистического моделирования [9]. При этом количество испытаний (объем выборок моделируемых распределений статистик), как правило, составлял величину $N = 10^6$.

III. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

В процессе исследований строились распределения статистик перечисленных критериев при справедливости гипотез H_i, $i = \overline{0,3}$, и различных объёмах выборок.

Характер зависимости распределений статистики B_n Фроцини при справедливости нулевой гипотезы иллюстрирует рисунок 1. Как видим, при ограниченных объемах выборок распределения статистики существенно зависят от n. Однако, при $n \geq 100$ распределения достаточно быстро сходятся к некоторому предельному закону.

Распределения статистик остальных критериев имеют более сильную зависимость от n. Поэтому в практических применениях приходится пользоваться таблицами процентных точек, приводимых для конкретных n. Рис.2 иллюстрирует зависимость от n распределений статистики T_1 (2) Хегази-Грина.

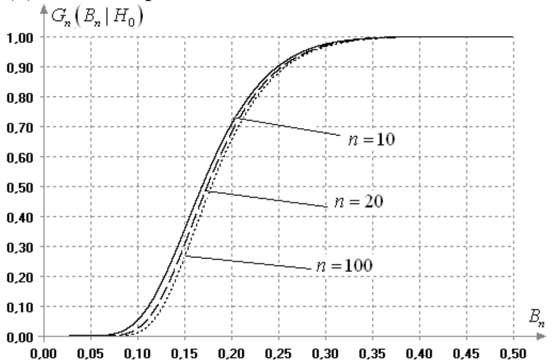

Рис. 1. Зависимость распределения $G(B_n | H_0)$ статистики B_n от объёма выборки

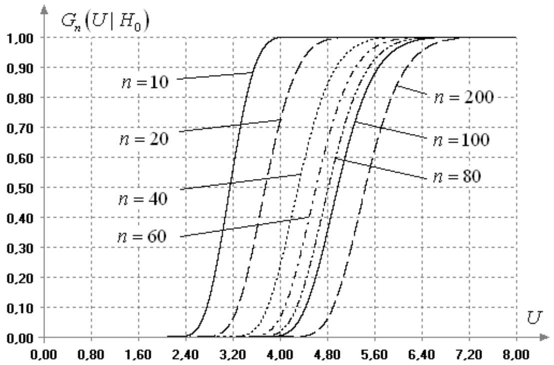

Рис. 2. Зависимость распределения $G(U | H_0)$ статистики U от объёма выборки

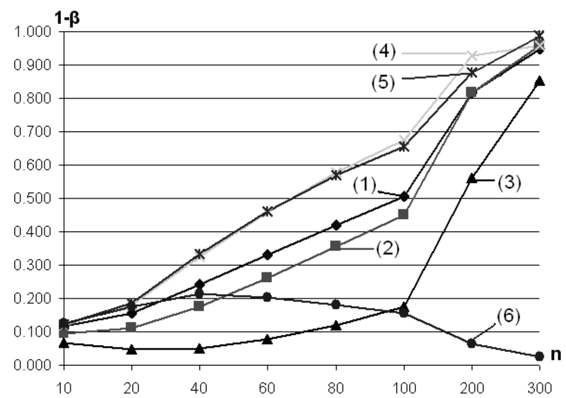

Рис. 3. Зависимость мощности критериев нормальности по отношению к конкурирующей гипотезе H_1 от объёмов выборок при $\alpha = 0.10$

Для каждого рассматриваемого объема выборки n (10, 20, 40, 60, 80, 100, 200, 300) для распределений статистик исследуемых критериев были вычислены таблицы процентных точек и найдены оценки мощности критериев по отношению к рассматриваемым конкурирующим гипотезам.

Полученные оценки мощности $1-\beta$ для каждого из критериев относительно гипотез H_1, H_2 и H_3 при уровне значимости $\alpha = 0.10$ в зависимости от n представлены на рисунках 3 – 5. Они позволяют судить о преимуществах в мощности того или иного критерия и подчеркивают смещённость некоторых критериев по отношению к некоторым конкурирующим гипотезам. Цифры на графиках соответствуют нумерации формул статистик критериев.

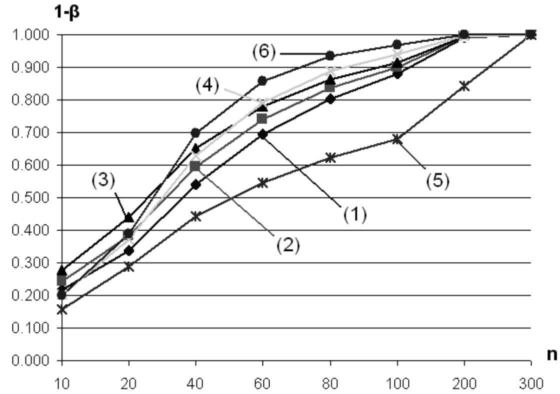

Рис. 4. Зависимость оценок мощности критериев нормальности по отношению к гипотезе H_2 от объёмов выборок при $\alpha = 0.10$

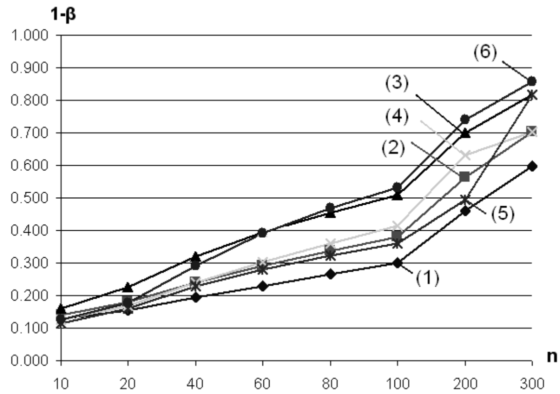

Рис. 5. Зависимость оценок мощности критериев нормальности по отношению к гипотезе H_3 от объёмов выборок при $\alpha = 0.10$

IV. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Исследуемые критерии можно разделить на две группы: несмещённые и смещённые относительно конкурирующей гипотезы H_1. Под смещённостью понимается то, что в определённых случаях мощность критерия меньше уровня значимости. Это значит, что при справедливости конкурирующей гипотезы критерий с большей уверенностью (!) будет считать выборку принадлежащей нормальному закону.

Критерий Хегази-Грина T_1 и T_2 являются смещёнными относительно гипотезы H_1 при $n \le 20$ и $n \le 60$ соответственно (при $\alpha \le 0.05$). Критерии Шапиро-Уилка и Эппса-Палли смещены относительно этой же конкурирующей гипотезы [2], но лишь при $n \le 20$ для $\alpha \le 0.05$.

Мощность критерия Шпигельхальтера по отношению к H_1 уменьшается с ростом $n > 40$. Но критерий является смещённым и при $n \ge 200$. Подобный недостаток, вообще говоря, нетипичен для критериев. По отношению к гипотезам H_2 и H_3 данный критерий сравним по мощности с критериями Шапиро-Уилка и Эппса-Палли при $n \le 20$ и заметно превосходит их при остальных n, Аналогичный вывод можно сделать, сравнивая его с критериями Фроцини и Хегази-Грина T_1 и T_2 при $n = 10$ и при остальных n соответственно.

Остальные четыре критерия не имеют смещения по отношению к рассматриваемым конкурирующим гипотезам.

V. ЗАКЛЮЧЕНИЕ

На основании анализа распределений статистик и оценок мощности критериев можно сделать следующие выводы.

Критерий Фроцини сравним по мощности с критериями Шапиро-Уилка и Эппса-Палли по отношению к гипотезам H_1 и H_2. По отношению к гипотезе H_1 уступает лишь критериям Гири и Дэвида–Хартли–Пирсона. Однако по отношению к H_2 превосходит лишь критерий Дэвида–Хартли–Пирсона, а в случае H_3 имеет наименьшую мощность.

Критерии Хегази-Грина T_1 и T_2 имеют достаточно высокую мощность по отношению к гипотезам H_2 и H_3 (их мощность выше мощности критериев Шапиро-Уилка и Эппса-Палли). Однако оба критерия являются смещенными в случае конкурирующей гипотезы H_1.

Критерий Дэвида-Хартли-Пирсона (5) по отношению к гипотезе H_1 превосходит по мощности критерии Шапиро-Уилка, Эппса-Палли и все остальные критерии, кроме критерия Гири. Но по отношению к H_2 и H_3 уступает критериям Шапиро-Уилка и Эппса-Палли, превосходя первый по отношению к H_3 лишь при $n \ge 40$.

Критерий Гири (4) является наиболее предпочтительным в случае конкурирующей гипотезы H_1 и мало уступает или не уступает остальным критериям относительно гипотез H_2 и H_3.

Недостатки критериев Шапиро-Уилка и Эппса-Палли, показанные в [2], могут компенсировать разумным применением других критериев, в частности, критериев Фроцини, Гири и Дэвида-Хартли-Пирсона, которые являются несмещёнными и могут использоваться при объёмах выборок $n \le 20$. При больших объёмах выборок целесообразно применять критерий Гири, как наиболее мощный из этих трёх.

Критерий Шпигельхальтера целесообразно применять, имея ввиду конкурирующие гипотезы типа логистического распределения или распределения Лапласа, но в случае конкурирующих гипотез близких к H_1 его использование бесполезно.

Работа выполнена при частичной финансовой поддержке РФФИ (проект № 06-01-0059а).

СПИСОК ЛИТЕРАТУРЫ

[1] ГОСТ Р ИСО 5479-2002. Статистические методы. Проверка отклонения распределения вероятностей от нормального распределения. - М.: Изд-во стандартов. 2002. - 30 с.

[2] Лемешко Б.Ю., Лемешко С.Б. Сравнительный анализ критериев проверки отклонения распределения от нормального закона // Метрология. 2005. №2. – С. 3-24.

[3] Frozini B.V. A survey of a class of goodness-of-fit statistics //Metron. – 1978. V.36. – № 1-2. – P.3-49.

[4] Frozini 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-New York: North Holland Publ.Comp, 1987, P.133-154.

[5] Hegazy Y.A.S., Green J.R. Some new goodness-of-fit tests using order statistics //Applied Statistics. – 1975. – V.24. – №3. – P.299-308.

[6] Spiegelhalter D.J. A test for normality against symmetric alternatives //Biometrika. – 1977. – V.64. – №2. – P.415-418.

[7] Geary R.C. The ratio of the mean deviation to the standard deviation as a test of normality // Biometrika. – 1935. – V.27. – P.310-322.

[8] David H.A., Hartley H.O., Pearson E.S. The distribution of the ratio, in a single normal sample, of range to standard deviation // Biometrika. – 1964. – V.512. - №3-4. – P.484-487.

[9] Лемешко Б.Ю., Постовалов С.Н. Компьютерные технологии анализа данных и исследования статистических закономерностей: Учеб. пособие. – Новосибирск: Изд-во НГТУ, 2004. – 120 с.

Андрей Павлович Рогожников — Бакалавр прикладной математики и информатики (2007 г.), магистрант факультета прикладной математики и информатики НГТУ.

Борис Юрьевич Лемешко — Профессор кафедры прикладной математики НГТУ, д.т.н., профессор.