

## ИССЛЕДОВАНИЕ РАСПРЕДЕЛЕНИЙ СТАТИСТИК И МОЩНОСТИ НЕПАРАМЕТРИЧЕСКИХ КРИТЕРИЕВ СОГЛАСИЯ JIN ZHANG ПРИ СЛОЖНЫХ ГИПОТЕЗАХ<sup>1</sup>

Козлова А.В., Лемешко Б.Ю.  
НГТУ, г. Новосибирск  
E-mail: kosaami@yandex.ru

Цель работы заключалась в исследовании предложенных в работах [1,2] непараметрических критериев согласия для ситуации проверки сложных гипотез. В частности, планировалось исследовать мощность новых критериев, сравнить с мощностью критериев типа Колмогорова, типа  $\omega^2$  Крамера-Мизеса-Смирнова и  $\Omega^2$  Андерсона-Дарлингга, исследовать распределения статистик данных критериев при проверке различных сложных гипотез и построить таблицы процентных точек при различных уровнях значимости и различных объемах выборок.

Применяя критерии согласия, различают проверку простых и сложных гипотез. Простая проверяемая гипотеза имеет вид  $H_0: F(x) = F_0(x, \theta)$ , где  $F_0(x, \theta)$  - функция распределения вероятностей, с которой проверяется согласие наблюдаемой выборки, а  $\theta$  - известное значение параметра (скалярного или векторного). Сложная проверяемая гипотеза имеет вид  $H_0: F(x) \in \{F_0(x, \theta), \theta \in \Theta\}$ . Будем предполагать, что оценка параметра распределения  $\hat{\theta}$  вычисляется по той же самой выборке, по которой проверяется согласие.

Статистики рассмотренных в работах [1], [2] непараметрических критериев согласия имеют вид:

$$Z_K = \max_{1 \leq i \leq n} \left( \left( i - \frac{1}{2} \right) \ln \left\{ \frac{i - \frac{1}{2}}{n F_0(X_{(i)})} \right\} + \left( n - i + \frac{1}{2} \right) \ln \left\{ \frac{i - \frac{1}{2}}{n [1 - F_0(X_{(i)})]} \right\} \right);$$
$$Z_A = - \sum_{i=1}^n \left[ \frac{\ln \{F_0(X_{(i)})\}}{n - i + \frac{1}{2}} + \frac{\ln \{1 - F_0(X_{(i)})\}}{i - \frac{1}{2}} \right];$$

<sup>1</sup> Работа выполнена при финансовой поддержке РФФИ (проект № 06-01-00059-а)

$$Z_C = \sum_{i=1}^n \left[ \ln \left\{ \frac{F_0(X_{(i)})^{-1} - 1}{\left(n - \frac{1}{2}\right) / \left(i - \frac{3}{4}\right) - 1} \right\} \right]^2.$$

Для принятия решения об отклонении или неотклонении проверяемой гипотезы  $H_0$  необходимо знать условное распределение  $G\{Z|H_0\}$  статистики  $Z$  соответствующего критерия или иметь таблицы процентных точек. В [1], [2] процентные точки приводятся для случая проверки простых гипотез и для проверки сложных гипотез о принадлежности нормальному закону.

В данной работе, используя методы статистического моделирования, были построены условные распределения статистик  $Z_A$ ,  $Z_C$ ,  $Z_K$  рассматриваемых критериев согласия при проверке простых гипотез и для случая, когда при проверке сложных гипотез параметры закона оценивались по этой же выборке методом максимального правдоподобия. По смоделированным выборкам значений статистик (по эмпирическим распределениям) оценивались процентные точки для различных уровней значимости.

Результаты исследований распределений статистик  $Z_A$ ,  $Z_C$ ,  $Z_K$  подтвердили, что при проверке простых гипотез критерии являются «свободными от распределения»: процентные точки распределений статистик критериев не зависят от вида закона, соответствующего  $H_0$ , с которым проверяется согласие. При проверке сложных гипотез свойство «свободы от распределения» теряется.

Недостатком критериев является то, что распределения статистик и значения процентных точек сильно зависят от объемов выборок  $n$ : с ростом  $n$  распределения статистик не сходятся к предельным распределениям.

В таблицах 1-2 приведены найденные значения мощности критериев для заданных вероятностей ошибок 1-го рода  $\alpha$  и различных объемов выборок  $n$  при проверке простых и сложных гипотез в случае, когда проверяемая гипотеза  $H_0$  соответствует нормальному закону, а конкурирующая гипотеза  $H_1$  – логистическому при значениях параметров, обеспечивающих наибольшую близость гипотез.

Сравнивая мощность данных критериев для рассматриваемой пары конкурирующих гипотез с мощностью критериев Колмогорова, Крамера-Мизеса-Смирнова, Андерсона-Дарлинга и критериев типа  $\chi^2$  [3], можно отметить следующее. В случае проверки простых гипотез они уступают по мощности лишь критериям типа  $\chi^2$  (в случае использования в последних асимптотически оптимального группирования). При проверке сложных гипотез они несколько превосходят по мощности критерии Колмогорова, Крамера-Мизеса-Смирнова, Андерсона-Дарлинга и критерии типа  $\chi^2$ . В таблицах 1 и 2 критерии упорядочены по убыванию мощности.

В данной работе для критериев со статистиками  $Z_A$ ,  $Z_C$ ,  $Z_K$  были построены таблицы процентных точек для проверки сложных гипотез о согласии с распределениями: экспоненциальным, полунормальным, Рэлея, Максвелла, Лапласа, нормальным, логнормальным, Коши, логистическим, максимального и минимального значения, Вейбулла. Эти результаты расширяют область применения данных критериев на практике. Недостатком новых критериев согласия является зависимость распределений статистик от объемов выборок  $n$  и отсутствие общей модели закона с параметрами, зависящими от  $n$ .

**Таблица 1.** Мощность критериев согласия при проверке простой гипотезы  $H_0$  против альтернативы  $H_1$

$\alpha$	$n=100$	$n=200$	$n=300$	$n=500$	$n=1000$	$n=2000$
	Мощность критерия $Z_C$					
0.1	0,313	0,475	0,585	0,751	0,950	0,999
0.05	0,259	0,387	0,495	0,664	0,917	0,999
0.01	0,155	0,258	0,342	0,517	0,812	0,989
	Мощность критерия $Z_A$					
0.1	0,264	0,372	0,486	0,680	0,933	0,999
0.05	0,163	0,238	0,335	0,520	0,863	0,999
0.01	0,044	0,069	0,126	0,228	0,635	0,970
	Мощность критерия $Z_K$					
0.1	0,233	0,326	0,427	0,603	0,868	0,994
0.05	0,146	0,217	0,297	0,460	0,749	0,975
0.01	0,047	0,079	0,118	0,233	0,467	0,866

**Таблица 2.** Мощность критериев согласия при проверке сложной гипотезы  $H_0$  против альтернативы  $H_1$

$\alpha$	$n=100$	$n=200$	$n=300$	$n=500$	$n=1000$	$n=2000$
	Мощность критерия $Z_C$					
0.1	0.361	0.540	0.668	0.839	0.981	1.000
0.05	0.304	0.464	0.596	0.786	0.969	1.000
0.01	0.186	0.328	0.456	0.667	0.928	0.999
	Мощность критерия $Z_A$					
0.1	0.351	0.511	0.639	0.819	0.978	1.000
0.05	0.257	0.391	0.518	0.725	0.954	1.000
0.01	0.113	0.191	0.282	0.481	0.846	0.996
	Мощность критерия $Z_K$					
0.1	0.350	0.499	0.616	0.779	0.954	0.999
0.05	0.249	0.379	0.490	0.667	0.906	0.996
0.01	0.110	0.188	0.267	0.421	0.731	0.968

1. Jin Zhang, Yuehua Wu. Likelihood-ratio tests for normality // *Computational Statistics & data analysis*. 49 (2005), – P.709-721.
2. Jin Zhang. Powerful goodness-of-fit tests based on the likelihood ratio // *Royal Statistical Society* (2002), Part 2. – P.281-294.
3. Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н. Мощность критериев согласия при близких альтернативах // *Измерительная техника*. 2007. № 2. – С.22-27.