

## О МОЩНОСТИ КРИТЕРИЕВ, ИСПОЛЬЗУЕМЫХ ДЛЯ ПРОВЕРКИ ГИПОТЕЗ О ПРИНАДЛЕЖНОСТИ ВЫБОРОК РАВНОМЕРНОМУ ЗАКОНУ

Блинов П.Ю., Лемешко Б.Ю.  
Новосибирский государственный технический университет  
E-mail: Blindizer@yandex.ru

Цель настоящей работы состояла в исследовании распределений статистик критериев, применяемых для проверки гипотез о принадлежности выборок равномерному закону, и в сравнительном анализе мощности совокупности статистических критериев.

Все рассмотренные критерии используют упорядоченную выборку (вариационный ряд), распределенную, как правило, на интервале  $[0,1]$ .

Статистики целого ряда критериев [1] предусматривают использование разностей последовательных значений вариационного ряда

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

где  $U_i$  –  $i$ -й элемент вариационного ряда,  $U_0 = 0$ ,  $U_{n+1} = 1$ ,  $n$  – объем выборки. К таким критериям относятся:

- критерий Шермана, статистика которого имеет вид

$$\omega_n = \frac{1}{2} \sum_{i=1}^{n+1} \left| D_i - \frac{1}{n+1} \right|;$$

- критерий Кимбелла со статистикой

$$w_k = \frac{1}{2} \sum_{i=1}^{n+1} \left( D_i - \frac{1}{n+1} \right)^2;$$

- критерий Морана со статистикой

$$M_n = - \sum_{i=1}^{n+1} \ln \left[ n+1 D_i \right];$$

- критерий Янга со статистикой

$$M = \frac{1}{x_n - x_1} \sum_{i=1}^{n-1} \min(D_i, D_{i+1});$$

- критерий Гринвуда-Кэсенберри-Миллера со статистикой

$$Q = \sum_{i=1}^{n+1} D_i^2 + \sum_{i=1}^n D_i * D_{i+1}.$$

Ряд критериев опирается на разности порядковых статистик, соответствующих анализируемой выборке, и математического ожидания этой порядковой статистики. К таким критериям можно отнести [1]:

- критерий Фроцини, статистика которого имеет вид

$$B_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left| U_i - \frac{i-0,5}{n} \right|;$$

- семейство критериев Хегази-Грина;
- критерии типа Колмогорова-Смирнова.

Критерий Неймана-Бартона опирается на отношение правдоподобия. В выражении для статистики используются величины

$$V_j = \frac{1}{\sqrt{n}} \sum_{i=1}^n \pi_j U_i,$$

где  $\pi_j U_i$  – полиномы Лежандра, ортогональные на интервале  $[0,1]$ . Применяются, как правило, первые 4 полинома. В общем случае статистика критерия имеет вид

$$N_K = \sum_{j=1}^K V_j^2,$$

где  $K$  – количество полиномов.

Статистика критерий Ченга-Спиринга имеет вид

$$W_p = \frac{\left[ U_n - U_1 \frac{n+1}{n-1} \right]^2}{\sum_{i=1}^n U_i - \bar{U}^2},$$

где разность  $U_n - U_1$  представляет собой выборочный размах.

При выполнении исследований опирались на методику компьютерного моделирования и исследования статистических закономерностей [2] и развиваемую программную систему [3].

При исследовании мощности критериев в качестве конкурирующих гипотез  $H_i$  (конкурирующих законов) рассматривались бета-распределения I рода  $B(\theta_0, \theta_1)$  с параметрами формы  $\theta_0$  и  $\theta_1$ , отличными от единицы, так как равномерное распределение является частным случаем  $B(1,1)$  этого закона.

Полученные в результате исследований оценки мощности при вероятности ошибки 1-го рода  $\alpha = 0.05$  и объемах выборок  $n = 50$  и  $n = 1000$  представлены в таблицах 1 и 2. В таблицах для соответствующих критериев введены следующие обозначения:  $\omega_n$  – критерий Шермана;  $M$  – Янга;  $M_n$  – Морана;  $W_p$  – Ченга-Спиринга;  $T_1, T_2, T_1^*, T_2^*$  – критерии Хегази-Грина и их модификации;  $D^+, D^-, D$  – модификации критериев типа Колмогорова-Смирнова;  $V$  – Купера;  $B_n$  – Фроцини;  $Q$  – Гринвуда-Кэсенберри-Миллера;  $N_2, N_3, N_4$  – Неймана-Бартона;  $w_k$  – Кимбелла.

**Выводы.** При исследовании распределений статистик критериев равномерности и оценке их мощностей было выделено несколько критериев, отличающихся большей эффективностью при проверке гипотезы о принадлежности выборок равномерному закону. При малых объемах выборок имеется преимущество у критериев Хегази-Грина и Неймана-

Бартона. В некоторых случаях хорошо показывают себя критерии типа Колмогорова-Смирнова и критерий Фроцини. При больших объемах выборок и конкурирующей гипотезе, близкой к равномерному распределению, хорошие результаты показывает критерий Ченга-Спиринга.

Таблица 1 – Мощности критериев при объемах выборок  $n = 50$

	Конкурирующие законы					Конкурирующие законы			
	B(0.9,0.9)	B(0.95,0.95)	B(1,2)	B(2,1)		B(0.9,0.9)	B(0.95,0.95)	B(1,2)	B(2,1)
$\omega_n$	0.065	0.056	0.529	0.531	$D^-$	0.061	0.054	0.994	0.000
$M_n$	0.053	0.051	0.286	0.287	$D$	0.063	0.055	0.983	0.983
$W_p$	0.075	0.056	0.360	0.361	$V$	0.084	0.062	0.879	0.879
$T_1$	0.063	0.056	0.994	0.994	$B_n$	0.060	0.054	0.995	0.995
$T_1^*$	0.057	0.053	0.995	0.996	$Q$	0.073	0.059	0.816	0.816
$T_2$	0.063	0.055	0.993	0.994	$N_2$	0.091	0.063	0.990	0.990
$T_2^*$	0.098	0.062	0.995	0.995	$N_3$	0.091	0.064	0.979	0.979
$M$	0.052	0.051	0.116	0.115	$N_4$	0.095	0.066	0.962	0.963
$D^+$	0.061	0.054	0.000	0.994	$w_k$	0.065	0.056	0.728	0.729

Таблица 2 – Мощности критериев при объемах выборок  $n = 1000$

	Конкурирующие законы				Конкурирующие законы		
	B(1.05,1.05)	B(1.1,1.1)	B(1.5,1.5)		B(1.05,1.05)	B(1.1,1.1)	B(1.5,1.5)
$\omega_n$	0.052	0.064	0.672	$D^-$	0.063	0.113	0.997
$M_n$	0.050	0.052	0.421	$D$	0.062	0.117	1.000
$W_p$	0.191	0.568	1.000	$V$	0.110	0.347	1.000
$T_1$	0.056	0.103	1.000	$B_n$	0.057	0.109	1.000
$T_1^*$	0.060	0.115	1.000	$Q$	0.056	0.084	0.995
$T_2$	0.060	0.106	1.000	$N_2$	0.148	0.480	1.000
$T_2^*$	0.062	0.115	1.000	$N_3$	0.123	0.413	1.000
$M$	0.051	0.052	0.072	$N_4$	0.120	0.415	1.000
$D^+$	0.063	0.116	0.997	$w_k$	0.054	0.072	0.962

Исследования наглядно показали, что среди множества рассмотренных не существует критерия равномерности, превосходящего другие критерии по всем аспектам. Этим и объясняется такое большое количество критериев, предложенных разными авторами.

### Литература

1. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников / А.И. Кобзарь. – М. : ФИЗМАТЛИТ, 2006. – 816 с.
2. 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. – 888 с.

3. **ISW** – Программная система статистического анализа одномерных случайных величин. URL: [http://www.ami.nstu.ru/~headrd/ISW\\_exe-r1006-2012.12.28-14\\_03.zip](http://www.ami.nstu.ru/~headrd/ISW_exe-r1006-2012.12.28-14_03.zip) (дата обращения 14.02.2013).