УДК 620.179

А.И. Федорович

ИССЛЕДОВАНИЕ НЕПАРАМЕТРИЧЕСКИХ КРИТЕРИЕВ СДВИГА В ЗАДАЧАХ НЕРАЗРУШАЮЩЕГО КОНТРОЛЯ

Аннотация. Проведен сравнительный анализ мощности непараметрических критериев сравнения сдвигов в выборках независимых и автокоррелированных случайных величин с различными законами распределения вероятностей. Исследована работоспособность критериев в задачах сравнения выборок с различными законами распределения вероятностей.

Ключевые слова: выборка измерений, мощность критерия, непараметрические критерии, сдвиг.

Постановка задачи

В задачах ультразвукового неразрушающего контроля объектов информация об их состоянии содержатся в выборках измерений, статистические закономерности которых, как правило, неизвестны. Изменение статистических закономерностей является свидетельством изменения состояния объекта. Оценка изменения может быть использована при контроле однотипных объектов и наблюдении за их состоянием при эксплуатации, после длительного хранения, перевозке, ремонте. Задача обнаружения изменения состояния объектов может быть решена путем сравнения двух выборок измерений с неизвестными законами распределения вероятностей. Чаще всего выборки отличаются центром группировки измерений и величиной рассеяния (сдвигом и масштабом по терминологии непараметрической статистики) [1]. В работе [1] рассмотрено десять методов использования непараметрических критериев сдвига для сравнения двух выборок случайных величин. Это критерии Кенуя, ранговый, Манна-Уитни-Вилкоксона, Ван-дер-Вардена, медианный, Мостеллера, Розенбаума, Хаги, Е-критерий. Критерии Мостеллера и Розенбаума имеют малоэффективны и поэтому не рассматриваются.

При сравнении выборок измерений возможны два ошибочных решения: 1) при равенстве сдвигов выборок случайных величин при-

ISSN 1562-9945

нимается противоположное решение (ошибки первого рода); 2) при сравнении двух выборок с неравными сдвигами принимаются ошибочные решения об их равенстве (ошибки 2-ого рода). Приведённые непараметрические критерии сдвига могут использоваться в задачах неразрушающего контроля для распознавания элементов с неизвестными случайными параметрами, измеряемые с ошибками и различными законами распределения вероятностей.

Цель исследования – оценка вероятностей ошибочных решений и сравнительный анализ работоспособности непараметрических критериев сдвига для различных видов законов распределения вероятности и коррелированности случайных величин в задачах неразрушающего контроля.

Вычислительные эксперименты

В работе рассмотрены критерии сдвига в контексте, что сравниваемые выборки случайных величин имеют одинаковый масштаб.

Исследуются выборки случайных величин с законами распределения вероятностей: логистическскским, Лапласа, Гаусса. И коррелированные нормальные случайные величины. Длина исследуемых выборок $n = 200$, а количество проведённых независимых экспериментов $k = 10000$. Эффективностью критерия является его способность распознавать различие между сдвигами исследуемых выборок. Вычислительные эксперименты проводились следующим образом:

1. Две выборки с логистическим (Лапласовским, нормальным) распределением с одинаковыми дисперсиями, математическое ожидание одной выборки фиксировано и равно нулю, а второй – изменяется от нуля до единицы, с шагом в 0,1.

2. Коррелированные нормальные случайные величины ($r = 0.1$, $r = 0.5$, $r = 0.9$), также при одном фиксированном математическом ожидании и изменяющимся другом, с тем же шагом.

3. Выборки с различными законами распределения вероятностей (логистический и нормальный; нормальный и Лапласа, логистический и Лапласа) фиксированным сдвигом одной выборки и изменяющимся – другой.

Анализ результатов вычислительных экспериментов

В результате проведения вычислительных экспериментов были получены таблицы эффективности непараметрических критериев сдвига. Пример такой таблицы приведен ниже, для двух выборок логистическим распределением ($\sigma 1 = \sigma 2 = 1$, $a1 = 0$). В таблице 1 приведено значение вероятности принятия правильного решения (сдвиги одинаковы) в зависимости от величины отклонения сдвига одной выборки от сдвига другой. При заданной вероятности принятия решения $P = 0.97$.

Таблица 1

№	Критерий \ $a2$	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
1	Стьюдента	0,963	0,837	0,482	0,135	0,018	0,003	0,001	0,001	0,001	0,001	
2	Ранговый	0,944	0,936	0,908	0,853	0,792	0,708	0,607	0,473	0,384	0,283	
3	Хаги	0,971	0,974	0,964	0,951	0,943	0,901	0,89	0,858	0,833	0,811	
4	Е	0,967	0,974	0,964	0,966	0,949	0,932	0,932	0,903	0,897	0,862	
5	Кенуя	0,706	0,7	0,707	0,696	0,688	0,682	0,682	0,663	0,663	0,646	
6	В-д-В	0,959	0,931	0,91	0,865	0,787	0,701	0,608	0,496	0,382	0,284	
7	Вилкоксона	0,956	0,942	0,927	0,848	0,8	0,71	0,596	0,503	0,383	0,293	
8	Медианный	0,927	0,916	0,88	0,831	0,773	0,722	0,612	0,505	0,434	0,336	

По результатам таблицы построены графики (рисунок 1).

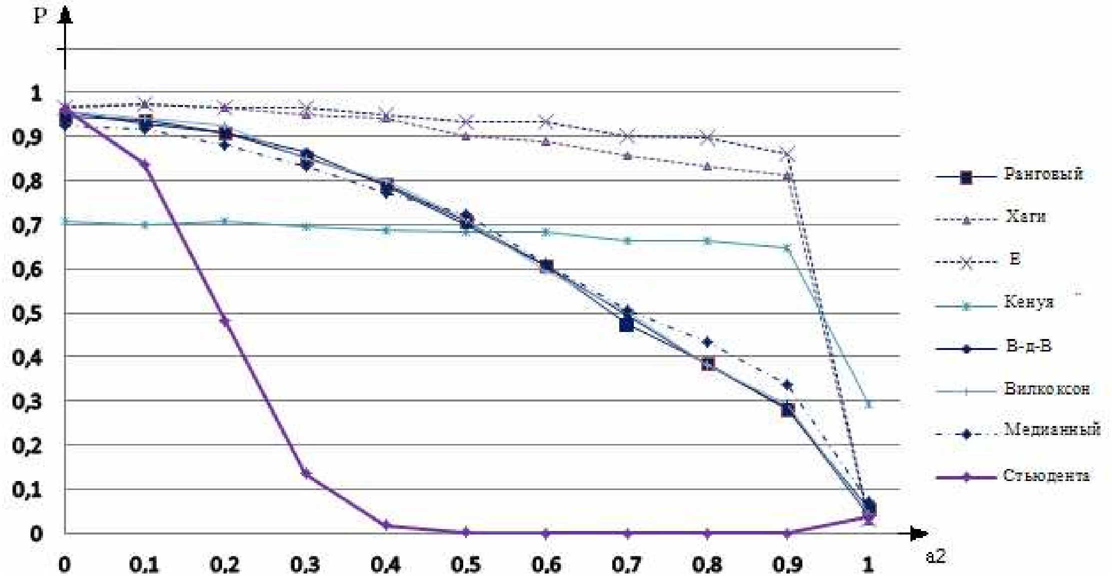

Рисунок 1 – Зависимость вероятности распознавания неравенства сдвигов от величины отклонения параметра $a2$ от $a1 = 0$

В таблице 2 приведён пример показателей эффективности непараметрических критериев сдвига при сравнении выборок с различными законами распределения вероятностей. Для случая сравнения выборок с нормальным и лапласовским распределениями ($\sigma 1 = \sigma 2 = 1$, $a1 = 0$). В таблице приведено значение вероятности, с которой каждый из критериев определяет выборки как имеющие различные сдвиги.

 ISSN 1562-9945

Таблица 2

№	$a2$ / Критерий	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
1	Стьюдент	0,046	0,078	0,186	0,386	0,615	0,802	0,926	0,974	0,993	0,996	1
2	Ранговый	0,088	0,083	0,156	0,281	0,425	0,589	0,705	0,825	0,907	0,958	0,982
3	Хаги	0,07	0,086	0,111	0,149	0,202	0,276	0,358	0,438	0,527	0,609	0,694
4	E	0,002	0,004	0,007	0,01	0,016	0,022	0,028	0,039	0,054	0,066	0,085
5	Кенуя	0,691	0,694	0,6989	0,68	0,669	0,653	0,64	0,624	0,603	0,579	0,553
6	В-д-В	0,04	0,062	0,125	0,241	0,369	0,573	0,717	0,842	0,92	0,976	0,992
7	Вилкоксон	0,039	0,065	0,135	0,272	0,429	0,603	0,766	0,877	0,949	0,978	0,995
8	Медианный	0,082	0,099	0,173	0,311	0,492	0,632	0,761	0,849	0,915	0,962	0,98

В случае коррелированных случайных величин, получаются следующие результаты. Приведён случай марковских случайных величин с $a1=1$ и коэффициентом корреляции $r=0.5$. Результаты данного эксперимента занесины в таблицу 3.

Таблица 3

№	$a2$ / Критерий	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
1	Стьюдент	0,247	0,538	0,886	0,987	1	1	1	1	1	1	1
2	Ранговый	0,254	0,33	0,513	0,718	0,866	0,942	0,984	0,996	1	1	1
3	Хаги	0,108	0,152	0,304	0,484	0,669	0,833	0,92	0,971	1	1	1
4	E	0,08	0,11	0,219	0,34	0,52	0,661	0,788	0,874	0,925	0,967	0,981
5	Кенуя	0	0	0	0	0	0	0	0	0	0	0
6	В-д-В	0,275	0,335	0,504	0,704	0,866	0,958	0,983	0,997	1	1	1
7	Вилкоксон	0,25	0,341	0,519	0,714	0,876	0,954	0,988	0,996	1	1	1
8	Медианный	0,22	0,272	0,507	0,645	0,798	0,922	1	1	1	1	1

Графики, построенные по результатам таблицы 3 приведены на рисунке 2.

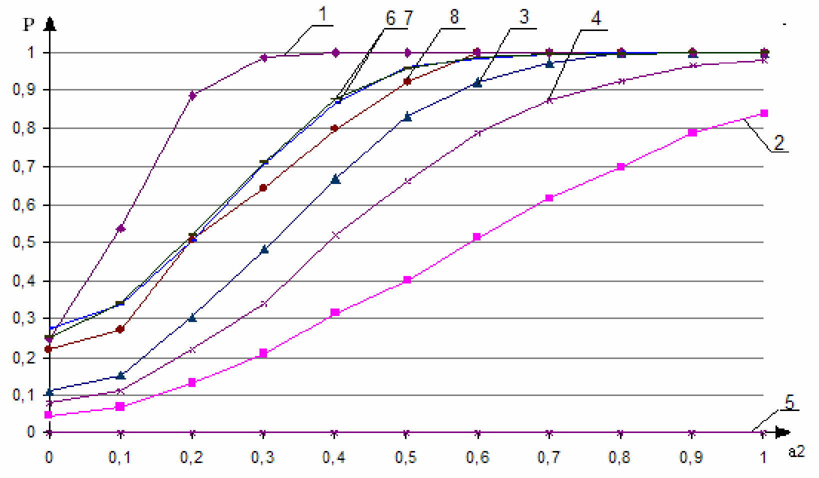

Рисунок 2 - Зависимость вероятности распознавания неравенства сдвигов от величины отклонения параметра $a2$ от $a1=0$, при корреляции выборок $r=0.5$

Из графиков и таблиц видно, что критерий Кенуя мало эффективен для любых из рассматриваемых случаев, и к применению не рекомендуется. Медианный критерий имеет невысокую эффективность распознавания различия между сдвигами (0.7) для любых сочетаний законов распределения выборок случайных величин. Неработоспособен для коррелированных случайных величин. Е-критерий имеет низкую эффективность распознавания при наличии изменения сдвигов, но может быть применён для выборок случайных величин с коэффициентов корреляции до 0.8. В этом случае критерий различает две выборки случайных величин, как имеющие различные сдвиги, с вероятностью 0.87. Критерий Хаги не работоспособен для коррелированных выборок. Но, эффективность не снижается, если исследуемые выборки случайных величин имеют различные законы распределения вероятностей. Ранговый критерий не применим для коррелированных случайных величин. Однако его эффективность (0.8-0.9) не снижается, если исследуемые выборки имеют различные законы распределения вероятностей. Критерий Ван-дер-Вардена и Вилкоксона идентичны по своим свойствам. Эффективность их распознавания наличия неравенства сдвигов двух выборок случайных величин составляет 0.8-0.9 для выборок с произвольными и даже различными видами законов распределения вероятностей. Но не применимы к выборкам коррелированным.

Поскольку, сама статистика каждого критерия является случайной величиной, то в дальнейшем будет исследовано влияние величины сдвигов на статистику каждого критерия, с целью определения по её параметрам, состояния объекта после эксплуатации.

Выводы

1. Непараметрические критерия сдвига возможно использовать лишь в тех случаях, когда заранее известно, что исследуемые выборки имеют одинаковые дисперсии.

2. Распознавание различия между сдвигами двух выборок случных величин не зависят от вида их закона распределения вероятностей. Наименее эффективным является критерий Кенуя.

3. Невозможно применение критериев для коррелированных случайных величин.

ЛИТЕРАТУРА

1. Кобзарь А.И. Прикладная математическая статистика/ А.И. Кобзарь. – М.: ФИЗМАТ ЛИТ, 2006. – 816 с.
2. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. – М.: Наука, 1983. – 416 с.

 ISSN 1562-9945