

А.И. Федорович

**ИССЛЕДОВАНИЕ НЕПАРАМЕТРИЧЕСКИХ КРИТЕРИЕВ
МАСШТАБА В ЗАДАЧАХ НЕРАЗРУШАЮЩЕГО КОНТРОЛЯ**

Аннотация. Проведен сравнительный анализ мощности непараметрических критериев сравнения масштабов в выборках независимых и автокоррелированных случайных величин с различными законами распределения вероятностей. Исследована работоспособность критериев в задачах сравнения выборок с различными законами распределения вероятностей. И статистика критериев масштаба во всех перечисленных случаях.

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

Постановка задачи

В задачах ультразвукового неразрушающего контроля объектов информация об их состоянии содержится 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 масштаба для сравнения двух выборок случайных величин. Это критерии Анасари-Бредли, Сижела-Тьюки, Кейпена, Клотца, квартильный критерий, Муда, Сэндвика-Олссона, Камата, Сукхатме, Сэвиджа и Краута-Линерта. Критерии Сукхатме, Сэвиджа и Краута-Линерта малоэффективны и поэтому здесь не рассматриваются.

Проведенные исследования в работе [2] изучены не все перечисленные критерии, и решалась задача только об обнаружении различия дисперсий в законах, отличных от нормального.

В задачах неразрушающего контроля при сравнении двух выборок измерений имеют места два ошибочных решения: 1) при равенстве масштабов выборок случайных величин принимается противоположное решение (ошибки первого рода); 2) при сравнении выборок с неравными масштабами принимаются ошибочные решения об их равенстве (ошибки 2-ого рода).

Цель исследования – анализ статистики непараметрических критериев масштаба при сдвиге масштабов, оценка вероятностей ошибочных решений 1-ого и 2-ого рода и работоспособности непараметрических критериев масштаба при различных видах законов распределения вероятности и коррелированности случайных величин. Эта задача решена путём проведения вычислительных экспериментов. Используя программы формирования выборок случайных величин с симметричными и асимметричными законами распределения вероятностей Лапласа и Гаусса, экспоненциальным и Релея, оценивалось влияние длины выборок и разностей их дисперсий, на законы распределения вероятностей исследуемых критериев и их параметры.

Эффективность критерия в задачах неразрушающего контроля – это его способность распознавать различие между масштабами исследуемых выборок. С этой целью решались следующие задачи:

1. Оценка критериев масштаба непараметрической статистики по величине минимальной ошибки первого рода.

2. Проведён факторный анализ исследуемых критериев путём оценивания влияния на статистические закономерности и их параметров видов законов распределения исследуемых выборок, их дисперсий, коррелированности и объёмов измерений (размеров выборок).

Анализ результатов вычислительных экспериментов

Первый вычислительный эксперимент имеет своей целью оценить эффективность всех критериев масштаба по величине вероятности ошибок первого рода при изменении дисперсий сравниваемых выборок. Исследовались две выборки по 100 измерений подчинённых нормальному закону распределения вероятностей с одинаковыми сдвигами и с изменяющейся дисперсией второй выборки (σ_2^2). Порог

сравнения критериев выбирался в соответствии с правилом Неймана-Пирсона (допустимая ошибка правильного решения P_{12} при отсутствии изменения масштаба выбрана не более 0,03). Результаты первого эксперимента представлены в таблице 1.

Таблица 1

σ_2/σ_1 Критерий	1	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	2
Анасари-Бредли	0,971	0,897	0,749	0,558	0,328	0,189	0,101	0,035	0,018	0,003	0,001
Сижела-Тьюки	0,97	0,96	0,881	0,823	0,783	0,6	0,521	0,433	0,315	0,21	0,193
Кейпена	0,937	0,855	0,868	0,811	0,79	0,698	0,626	0,599	0,57	0,507	0,461
Клотца	0,986	0,948	0,725	0,522	0,26	0,102	0,024	0,002	0	0	0
Квартильный	0,953	0,936	0,846	0,709	0,557	0,392	0,294	0,181	0,105	0,055	0,03
Муда	0,953	0,882	0,717	0,47	0,248	0,141	0,042	0	0	0	0
Сэндвика-Олссона	0,951	0,897	0,717	0,508	0,286	0,139	0,049	0,016	0,003	0	0
Камата	0,951	0,891	0,743	0,549	0,347	0,233	0,13	0,058	0,028	0,009	0

Из её анализа следует, что лучшими критериями сравнения выборок является критерий Клотца ($P_{12}^* = 0,14$) и критерий Анасари-Бредли ($P_{12}^* = 0,29$). Они же лучше и при обнаружении изменений масштаба. Так, если $\sigma_2/\sigma_1 = 1$, то вероятность обнаружения этого изменения критерием Клотца $P_{22}^* = 0,976$ ($P_{12}^* = 0,024$) и критерием Анасари-Бредли $P_{22}^* = 0,899$ ($P_{12}^* = 0,101$).

Второй эксперимент – оценка влияния на ошибочные решения размеров выборок измерений при использовании критериев Клотца и Анасари-Бредли. Результаты эксперимента представлены в таблице 2.

Таблица 2

n Критерий	10	20	30	50
Анасари-Бредли	0,035	0,029	0,024	0,021
Клотца	0,018	0,016	0,015	0,014

Очевидно, что с увеличением количества измерений ошибки уменьшаются, однако рассматриваемые критерии позволяют получить достаточно точные результаты и при незначительных объёмах

исходных данных, что расширяет возможности их использования и сокращает расходы по накоплению данных.

Исследовались выборки с одинаковыми параметрами, но различные законы распределения вероятностей: нормальной и/или экспоненциальный. Установлено как в этих условиях изменяется статистика критериев и их параметры. Результаты эксперимента представлены в виде гистограмм и таблиц. На рисунке 1 приведены гистограммы критерия Клотца для случаев входных выборок с нормальным и/или экспоненциальным законами распределения вероятностей при длинах выборок $n = 50$ и одинаковых значениях параметров.

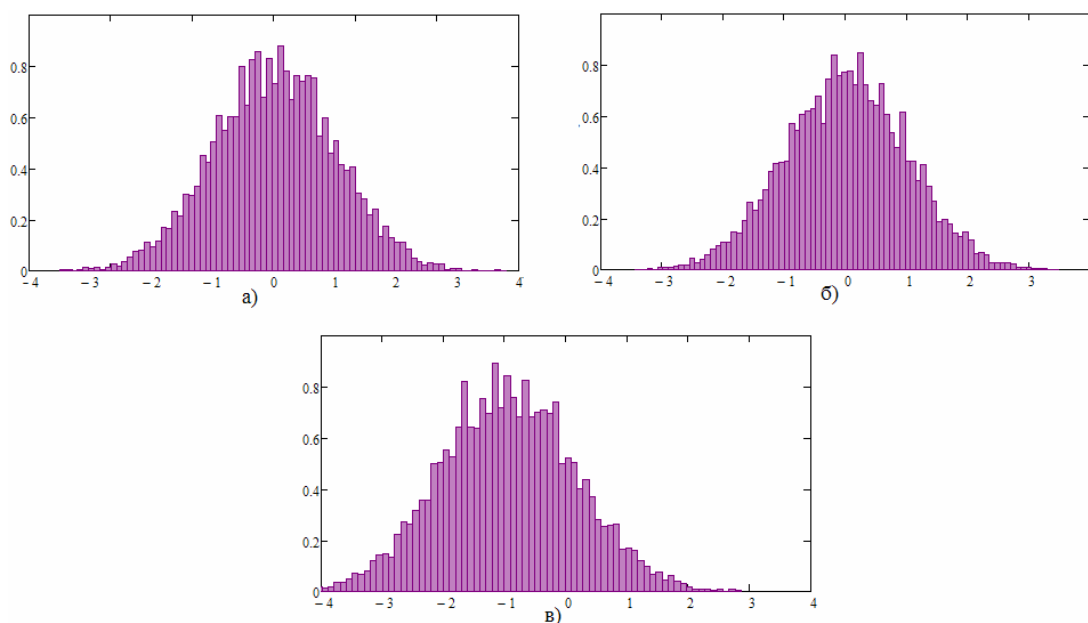


Рисунок 1 – Гистограммы статистики критерия Клотца

а) выборки с нормальным законом ($a_1 = a_2 = 1, \sigma_1 = \sigma_2 = 1$);

б) выборки с экспоненциальным законом ($\lambda_1 = \lambda_2 = 1$); в) выборки с

нормальным ($a_1 = 1, \sigma_1 = 1$) и экспоненциальным законом ($\lambda_2 = 1$)

Из рисунка 1 видно, что статистика критерия подчиняется нормальному распределению вероятностей. Эта гипотеза проверена по критерию хи-квадрат, и получила подтверждение (показатель критерия хи-квадрат $z = 110,542$, при пороговом значении $z_0 = 124,34$, то есть выполняется условие $z < z_0$). Это нормальное распределение имеет нулевое математическое ожидание и единичную дисперсию. Исключение составляет случай, когда сравниваются выборки с нормальным и экспоненциальным распределениями вероятностей, тогда математическое ожидание статистики становится равным -1. Такая закономер-

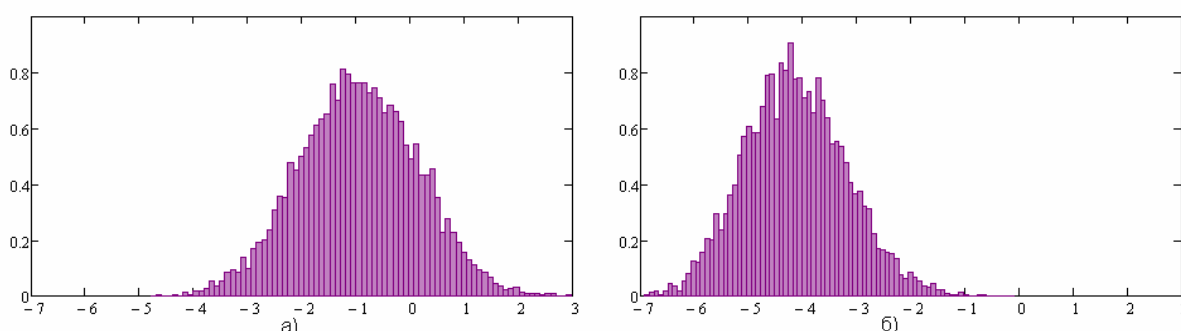
ность справедлива и для прочих вариантов различных законов распределения вероятностей при равных параметрах. Аналогичный эксперимент для критерия Анасари-Бредли подтвердил, что его статистика так же подчиняется нормальному распределению.

В таблице 3 приведены показатели эффективности непараметрических критериев масштаба при сравнении выборок с нормальным ($\mu_1 = 1 \sigma_1 = 1$) и экспоненциальным ($\lambda_2 = 1$) распределениями. В таблице приведено значение вероятности, с которой каждый критерий определяет выборки как имеющие различные масштабы ($n = 50$).

Таблица 3

σ_2	1	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	2
Критерий											
Анасари-Бредли	0,519	0,325	0,229	0,135	0,097	0,035	0,024	0,01	0	0	0
Клотца	0,545	0,274	0,102	0,031	0,004	0	0	0	0	0	0

При этом статистика критериев не изменяется, но различие масштабов оказывает существенное влияние на математическое ожидание распределения критериев. На рисунке 2 представлены гистограммы статистики критерия Клотца для выборок с нормальным и экспоненциальным распределением вероятностей. Рисунок 2а – масштабы выборок совпадают и равны 1 ($n = 50$), рисунок 2б - $\sigma_2/\sigma_1 = 2$.

Рисунок 2 – Гистограммы статистики критерия Клотца а) $\sigma_2/\sigma_1 = 1$,

$$б) \sigma_2/\sigma_1 = 2$$

Поскольку параметры статистики самих критериев зависят от разности масштабов исходных данных, то этот факт можно использовать для определения неоднородности выборок измерений.

Исследовалось влияние корреляции на эффективность непараметрических критериев масштаба. Приведён случай коррелированных ($r=0.9$) нормальных случайных величин с нулевым математическим ожиданием при фиксированном значении одной дисперсии $\sigma_1^2=1$ изменяющимся значении σ_2^2 . Объем экспериментальных выборок $n=50$. Результаты эксперимента занесены в таблицу 3.

Таблица 3

σ_2 / σ_1 Критерий	1	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	2
Анасари-Бредли	0,841	0,777	0,72	0,644	0,61	0,544	0,484	0,435	0,392	0,357	0,311
Клотца	0,693	0,68	0,629	0,566	0,453	0,371	0,322	0,276	0,196	0,162	0,147

Из графиков и таблиц видно, что критерий Анасари-Бредли обладает высокой эффективностью распознавания различия масштабов (0,96) и не теряет своих свойств распознавания даже для коррелированных выборок, с коэффициентом корреляции не превышающим 0,7. Критерий Клотца обладает высокой эффективностью. Распознаёт разницу в масштабах 0,1 с величиной ошибки первого рода до 0,08. Но, не пригоден для коррелированных выборок, или выборок с различными законами распределения вероятностей.

Выводы

1. Непараметрические критерии масштаба целесообразно использовать только в тех случаях, когда заранее известно, что изменяется только параметр масштаба. А исследуемые выборки имеют одинаковое математическое ожидание, но величина его не влияет на эффективность работоспособность критериев.

2. В результате проведения вычислительных экспериментов установлено, что в случае корреляции исследуемых выборок непараметрические критерии масштаба теряют эффективность.

3. Критерии масштаба являются случайной величиной и моделью их статистических закономерностей может служить нормальный закон распределения при сравнении как однородных, так и неоднородных выборок измерений.

4. Для неоднородных выборок величина математического ожидания нормального закона (для критериев Анасари-Бредли, Камата,

Сижела-Тьюки и Кейпена) линейно зависит от степени разности масштабов исходных выборок.

ЛИТЕРАТУРА

1. Кобзарь А.И. Прикладная математическая статистика/ А.И. Кобзарь. – М.: ФИЗМАТ ЛИТ, 2006. – 816 с.
2. Горбунова А.А., Лемешко Б.Ю., Лемешко С.Б. Критерии проверки гипотез об однородности дисперсий при наблюдаемых законах, отличных от нормальны. – Материала X международной конференции «Актуальный проблем электронного приборостроения» АПЭП – 2010. Т.6, Новосибирск, 2010. – с. 36-41.