

УДК 620.179

А.И. Федорович

КРИТЕРИЙ БУША-ВИНДА В ЗАДАЧАХ МОНИТОРИНГА ТЕХНИЧЕСКИХ ОБЪЕКТОВ

Аннотация. Исследуются изменения статистических закономерностей критерия непараметрической статистики Буша-Винда, чувствительность критерия к изменениям сдвига и масштаба выборок, содержащих информацию о состоянии технических объектов, контроль и мониторинг которых проводится в условиях полного незнания законов распределения вероятности и ограничений на объем измерений.

Ключевые слова: выборка измерений, непараметрическая статистика, критерий, вероятностные характеристики, сдвиг, масштаб.

Постановка задачи

Информация о состоянии технических линейно-протяженных объектов контроля содержится в выборках измерений параметров, которые изменяются во времени, и их статистические закономерности неизвестны. В задачах наблюдения за их состоянием (мониторинга) объекты периодически контролируются, выборки измерений сравниваются и по результатам сравнения должны приниматься решения о степени опасности изменений состояния объектов. В работе [1] показано, что для этих целей можно использовать критерии непараметрической статистики сдвига и масштаба, в частности критерий Буша-Винда. Его статистические закономерности известны при сравнении двух однородных выборок и неизвестны, если в выборках изменились параметры или законы распределения вероятности. Учитывая трудности аналитического исследования статистических закономерностей при сравнении неоднородных выборок, эта задача решалась на основе результатов вычислительных экспериментов путем моделирования и сравнения выборок случайных величин с различными законами распределения вероятности и различными параметрами. Цель исследований – оценить статистические закономерности критерия Буша-Винда при сравнении неоднородных выборок.

© Федорович А.И., 2010

Критерии Буша-Винда в задачах проверки гипотез о равенстве параметров сдвига и масштаба.

Критерий Буша-Винда, предназначенный для проверки гипотезы о равенстве неизвестных математических ожиданий и дисперсий двух выборок измерений, является комбинацией аналогов критериев Ван-дер-Вардена S и Клотца T

$$S = \sqrt{2\left(2 - \frac{1}{n}\right)} \frac{\sum_{i=1}^n \Psi \left[\frac{R(x_{2i})}{2n+1} \right]}{\sqrt{\sum_{i=1}^{2n} \Psi^2 \left[\frac{i}{n+1} \right]}},$$

$$T = \sqrt{2\left(2 - \frac{1}{n}\right)} \frac{\sum_{i=1}^n \Psi^2 \left(\frac{R(x_{2i})}{2n+1} \right) - \sum_{i=1}^{2n} \Psi^2 \left(\frac{i}{2n+1} \right)}{\sqrt{\sum_{i=1}^{2n} \Psi^4 \left(\frac{i}{2n+1} \right)}}, \quad (1)$$

где $R(x_{2i})$ - ранг i -ого измерения x_{2i} второй выборки в объединенной выборке измерений $R(x_{2i}) = \sum_{j=1}^n \text{sign}(x_{2i} - \xi_j)$, $\xi_1 > \xi_2 > \dots > \xi_{2n-1} > \xi_{2n}$.

Критерий Буша-Винда W равен комбинации критериев S и T вида

$$W = -2 \ln \left[2(1 - \Phi(|S|)) - 2 \ln \left[2(1 - \Phi(|T|)) \right] \right],$$

где $\Phi(x)$ - интеграл вероятности Гаусса, для вычисления которого можно применить аппроксимацию

$$\Phi^*(x) = 1 - 0.852 \exp \left[- \left(\frac{x + 1,5774}{2,0637} \right)^{2,34} \right], \quad x \geq 0, \quad \Phi^*(-x) = 1 - \Phi^*(x), \quad (2)$$

$\Psi(z)$ - функция, обратная интегралу вероятности Гаусса, для которой можно применить аппроксимацию вида

$$\Psi(z) = 4,91 \left(z^{0,14} - (1-z)^{0,14} \right) \quad (3)$$

Показатели Буша-Винда W являются случайными величинами и при $n > 30$ справедлива модель их статистических закономерностей законом распределения хи-квадрат с четырьмя степенями свободы, если законы распределения, математические ожидания и дисперсии

двух сравниваемых выборок равны. В этом случае с вероятностью $P = 0,95$ будет выполняться неравенство $W \leq 9,5$.

Для проведения исследований была разработана компьютерная модель формирования выборок измерений с различными законами распределения вероятностей с заданными параметрами. Путем сравнения этих выборок оценивалась статистика критерия Буша-Винда. Моделировались выборки случайных величин с симметричными, слабо и сильно асимметричными законами распределения Гаусса, Релея и экспоненциальным. Изменялись размеры выборок, их параметры, исследовались гистограммы критерия Буша-Винда, оценивались математические ожидания и дисперсии.

Статистика критерия Буша-Винда при сравнении неоднородных выборок.

По результатам вычислительных экспериментов установлено, что критерий Буша-Винда является критерием однородности сравниваемых выборок случайных величин, так как его статистика меняется не только при измерении сдвига и масштаба, но и при изменении только вида закона распределения при нулевом сдвиге и масштабе [1]. Следовательно, критерий Буша-Винда можно использовать в задачах оценки неоднородности двух выборок без знания их статистических закономерностей.

Известно, что если две исследуемые выборки случайных величин имеют одинаковые законы распределения вероятностей и одинаковые параметры сдвига и масштаба, то независимо от вида этих законов критерий Буша-Винда как случайная величина имеет закон распределения хи-квадрат с четырьмя степенями свободы. Гистограммы показателя Буша-Винда при сравнении однородных (а) и неоднородных (б) со сдвигом $\frac{\Delta a}{\sigma_1} = 0,8$ выборок с нормальными законами распределения вероятностей при единичных дисперсиях показаны на рис.1.

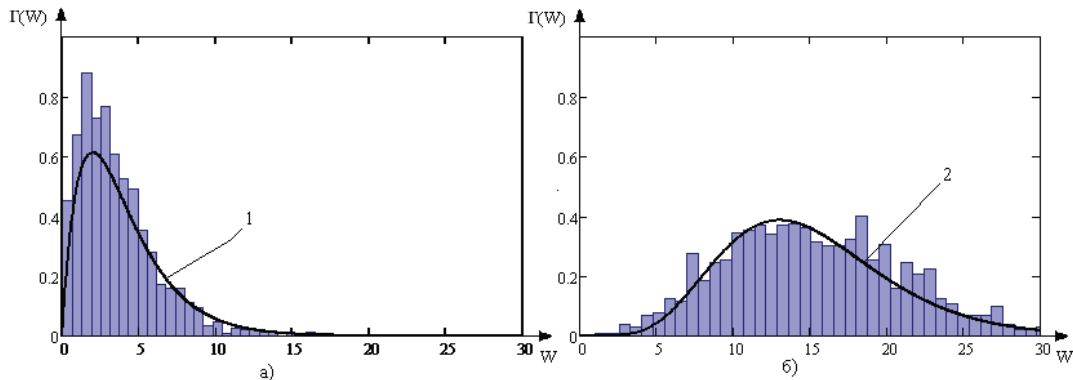


Рисунок 1 – Гистограммы показателя Буша-Винда ($n = 30$)

На рис. 2 представлены гистограммы показателя Буша-Винда при сравнении нормальных выборок с нулевым сдвигом и изменяющимися дисперсиями (а), и с изменением как математического ожидания $\Delta a / \sigma_1 = 0,8$, так и дисперсии $\Delta \sigma / \sigma_1 = 0,8$ (б).

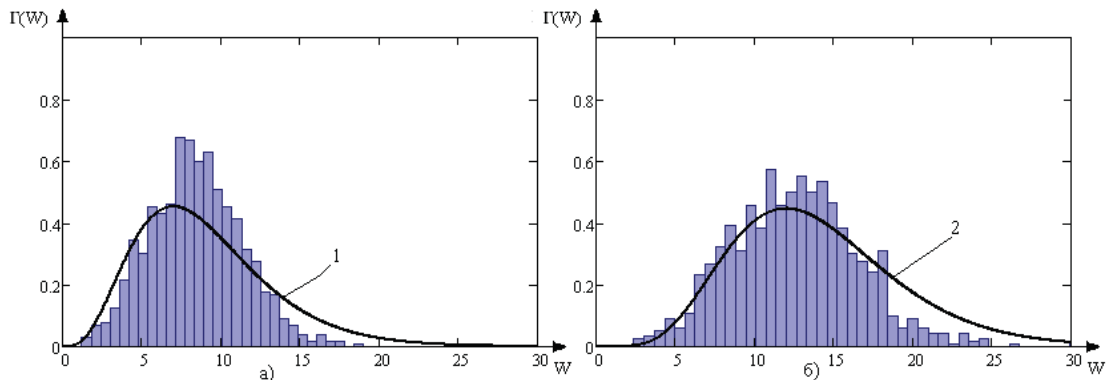


Рисунок 2 - Гистограммы критерия Буша-Винда

На рис. 3 показаны гистограммы показателя Буша-Винда для выборок с экспоненциальным (3а) и релеевским (3б) распределением.

С изменением сдвигов $\Delta a / \sigma_1 = \Delta \lambda / \lambda_1 = 0,8$;

$\Delta a / \sigma_1 = \sqrt{\frac{\pi}{4 - \pi}} \left(\sqrt{\frac{b_2}{b_1}} - 1 \right) = 0,654$, и масштабов $\Delta \sigma / \sigma_1 = \Delta \lambda / \lambda_1 = 0,8$;

$\Delta \sigma / \sigma_1 = \sqrt{\frac{b_1}{b_2}} = 0,745$.

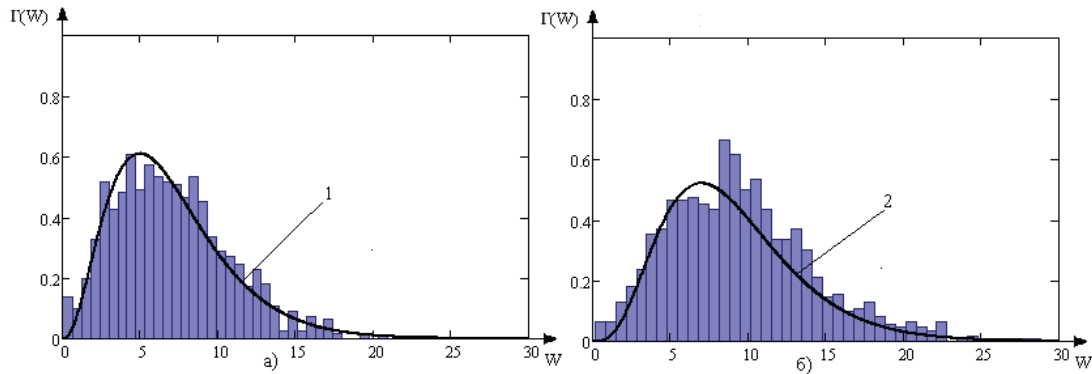


Рисунок 3. – Гистограммы показателя Буша-Винда для выборок с экспоненциальным (а) и релеевским (б) распределением

На рис.1,2,3 показаны графики закона распределения вероятности хи-квадрат для однородных выборок (рис.1а) и неоднородных выборок (рис. 1б, 2, 3). Параметр закона распределения вероятности хи-квадрат (степень свободы N) выбран по максимуму гистограммы ($W_{\max} = N - 2$). В данном случае число степеней свободы на рис.1а равно 4, $N=12$ (рис.1б); $N=8$ (рис.2а); $N=13$ (рис.2б); $N=6$ (рис.3а); $N=8$ (рис.3б). Гипотеза о виде закона распределения вероятности критерия W проверялась по критерию хи-квадрат. Показатель хи-квадрат для выборок с нормальным законом $z = 38,97$, с релеевским - $z = 64,652$, с экспоненциальным - $z = 47,843$ при пороговом значении $z_0 = 70,952$.

Проведен анализ зависимости средних значений $M^*[W] = \bar{W}$, и средне квадратичных отклонений $\sqrt{D^*[W]}$ критерия Буша-Винда при сравнении выборок с различными степенями неоднородности. В таблице 1 приведены эти данные для нормальных выборок размером $n = 30$, полученные путем обработки результатов вычислительных экспериментов (1000 реализаций).

Для однородных выборок математическое ожидание $M[W]$ и дисперсия $D[W]$ показателя Буша-Винда при $n > 30$ равны 4 и 8 соответственно. По результатам экспериментов оценки математических ожиданий и среднеквадратичные отклонения несколько меньше теоретических значений (4 и 2,83).

При изменении сдвига в 2 раза среднее значение показателя \bar{W} увеличилось с 3,68 до 19,21 (примерно в 5 раз),

среднеквадратичное значение изменилось примерно в 1,8 раза. При увеличении масштаба в 2 раза, среднее значение изменяется с 3,62 до 9,46 (в 3 раза), среднеквадратическое значение практически не изменилось (2,58 и 2,89). При одновременном изменении сдвига и масштаба в 2 раза эти показатели увеличились в 3 и 1,5 раза.

Таблица 1

$\Delta a / \sigma_1$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
\bar{W}	3,68	4,12	5,12	6,43	7,64	9,37	11,02	13,12	15,42	17,32	19,21
$\sqrt{D^* [W]}$	2,50	2,81	3,52	4,04	4,51	4,85	5,23	5,64	5,95	6,19	6,4
$\Delta \sigma / \sigma_1$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
\bar{W}	3,62	3,79	1,30	4,98	5,62	6,28	6,98	7,78	8,27	8,90	9,46
$\sqrt{D^* [W]}$	2,58	2,58	2,67	2,78	2,81	2,83	2,86	2,97	2,87	2,84	2,89
$\Delta a / \sigma_1$ $\Delta \sigma / \sigma_1$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
\bar{W}	3,88	4,29	4,81	5,65	6,40	7,34	8,31	9,16	9,87	10,76	14,71
$\sqrt{D^* [W]}$	2,63	2,90	2,93	3,27	3,31	3,57	3,65	3,84	3,59	3,77	3,91

В таблицах 2 и 3 приведены результаты вычислительных экспериментов для экспоненциального и релейского распределения вероятности соответственно, из которых следует, что среднее значение и среднеквадратичное отклонение показателя Буша-Винда примерно так же как и для нормальных выборок.

Таблица 2

$\Delta a / \sigma_1$ $\Delta \sigma / \sigma_1$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
\bar{W}	3.6 7	3.5 9	3.7 7	4.01	4.51	5.06	5.60	6.11	6.87	7.55	7.9 9
$\sqrt{D [W]}$	2.4 0	2.5 2	2.3 6	2.56 5	2.78 5	3.10 2	3.30 8	3.39 7	3.49 3	3.69 2	4.0 3

Таблица 3

$\Delta a / \sigma_1$ $\Delta \sigma / \sigma_1$	0	0.093	0.183	0.268	0.351	0.43	0.507	0.581	0.724	0.792
\bar{W}	3,61	4,11	4,75	5,47	6,087	6,991	7,94	8,667	10,32	11,37
$\sqrt{D [W]}$	2,49	2,81	3,03	3,753	3,877	4,106	4,19	4,492	4,872	4,882

Поскольку релеевское и экспоненциальное распределения являются однопараметрическими, то при изменении значения параметра в 2 раза одновременно изменятся сдвиг и масштаб. Для экспоненциального распределения значение показателя и дисперсии также увеличивается в 2 раза. Для релеевского распределения значения показателя изменяется в 3 раза, а его дисперсия в 2 раза.

Чувствительности критерия Буша-Винда к изменению параметров сравниваемых выборок

Путем проведения вычислительных экспериментов исследовалась чувствительность критерия как зависимость среднего значения показателя \bar{W} от изменения сдвига и масштаба для выборок различной длины (n). На рис. 4 представлены зависимости среднего значения критерия Буша-Винда \bar{W} от сдвига для выборок с нормальным распределением. Зависимость показателя Буша-Винда от изменения сдвига можно описать линейной моделью вида $\bar{W} = a + b \frac{\Delta a}{\sigma_1}$ при $\frac{\Delta a}{\sigma_1} > 0.2$. Значения коэффициентов a, b приведены в таблице 4.

Таблица 4

n	10	20	30	50	100
a	4,199	2,411	2,841	1,623	1,276
b	0,345	4,498	7,31	12,5	24,982

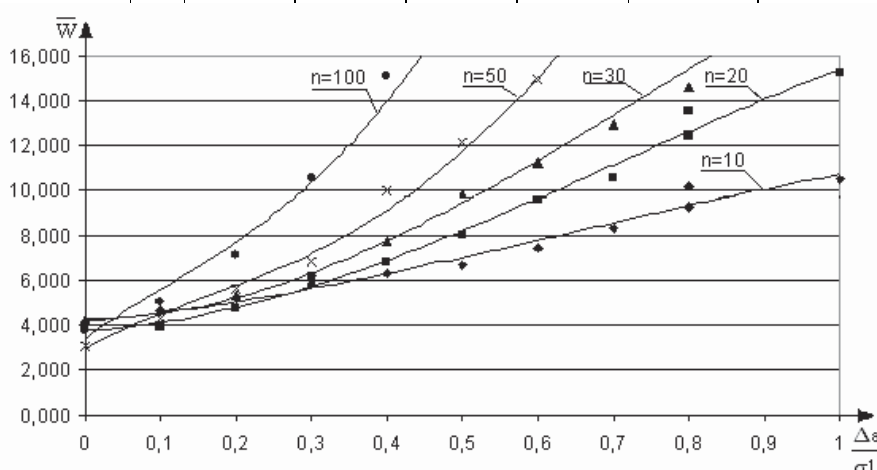


Рисунок 4 – Кривые чувствительности критерия Буша-Винда для выборок с нормальным распределением при изменении сдвига

Такие же эксперименты были проведены для нормальных выборок при изменении масштаба. Зависимости полученные в результате этих экспериментов, приведены на рис.5.

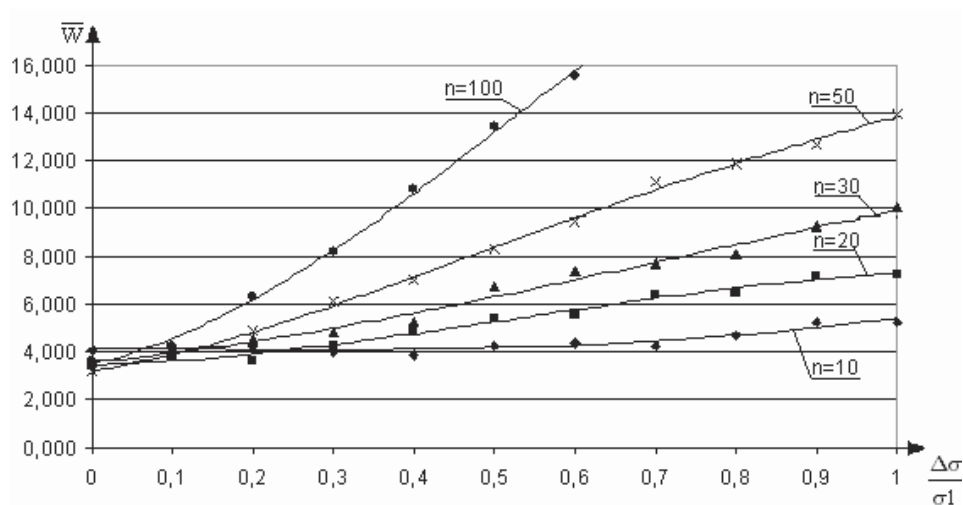


Рисунок 5 – Кривые чувствительности критерия Буша-Винда для выборок с нормальным распределением при изменении масштаба

Критерий Буша-Винда гораздо чувствительней к изменениям сдвига, чем к изменениям масштаба. При изменении сдвига даже на коротких выборках ($n = 10$) наблюдается превышение показателем \bar{W} порогового значения W_0 .

На рис. 6 приведены зависимости чувствительности при одновременном изменении сдвига и масштаба нормальных случайных величин.

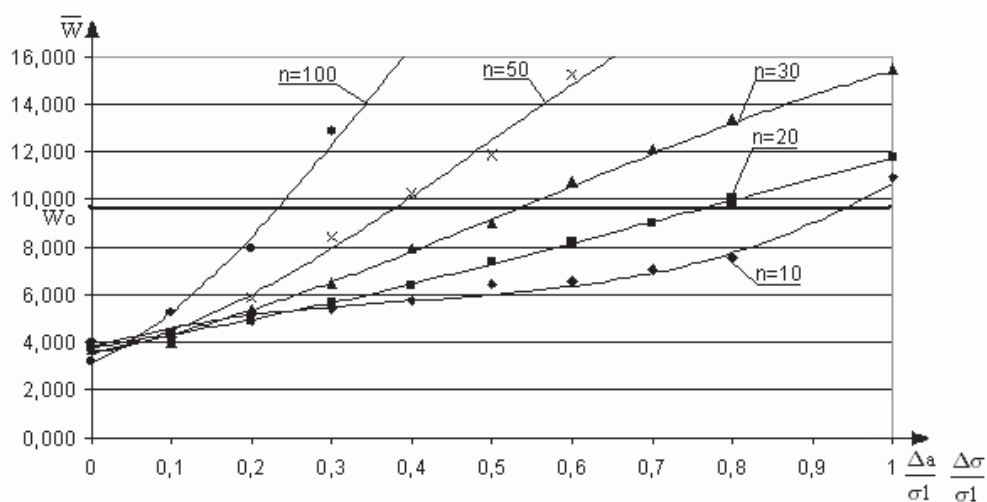


Рисунок 6 – Кривые чувствительности критерия Буша-Винда для выборок с нормальным распределением при изменении сдвига и масштаба

Аналогичные данные были получены при вычислительных экспериментах с выборками других симметричных и асимметричных

законов распределения вероятностей. Установлено, что чувствительность критерия непараметрической статистики Буша-Винда зависит не только от объема выборок измерений, но и от вида закона распределения вероятности.

Выводы

1. Критерий Буша-Винда как показатель однородности выборок случайных величин может использоваться для оценки степени неоднородности измерений при неизвестных законах распределения вероятности и их параметров в задачах мониторинга технических объектов.

2. Критерий Буша-Винда является случайной величиной и моделью его статистических закономерностей может служить закон распределения хи-квадрат при сравнении как однородных, так и неоднородных выборок измерений.

3. Для неоднородных выборок число степеней свободы закона хи-квадрат зависит от размеров сравниваемых выборок и степени различия их параметров, и его значение примерно равно среднему значению \bar{W} .

ЛИТЕРАТУРА

1. Малайчук В.П., Лысенко Н.А., Федорович А.И. Критерии непараметрической статистики Клотца и Буша-Винда в задачах периодического контроля технических объектов.// Системные технологии. Региональный межвузовский сборник научных работ. – Выпуск 2 (67) - Днепропетровск, 2010.-с.198-205.
2. Кобзарь А.И. Прикладная математическая статистика/ А.И. Кобзарь. – М.: ФИЗМАТ ЛИТ, 2006. – 816 с.
3. Малайчук В.П., Мозговой А.В. Математическая дефектоскопия: Монография. – Днепропетровск: Системные технологии, 2005, - 180 с.

Получено 09.04.2010г.