СЕДЬМАЯ МЕЖДУНАРОДНАЯ
НАУЧНО-ТЕХНИЧЕСКАЯ КОНФЕРЕНЦИЯ

ИЗМЕРЕНИЯ И ИСПЫТАНИЯ В СУДОСТРОЕНИИ И СМЕЖНЫХ ОТРАСЛЯХ

СУДОМЕТРИКА 2018

МАТЕРИАЛЫ КОНФЕРЕНЦИИ

15-17 ОКТЯБРЯ 2018 г.
САНКТ-ПЕТЕРБУРГ

СЕДЬМАЯ МЕЖДУНАРОДНАЯ
НАУЧНО-ТЕХНИЧЕСКАЯ КОНФЕРЕНЦИЯ

ИЗМЕРЕНИЯ И ИСПЫТАНИЯ В СУДОСТРОЕНИИ И СМЕЖНЫХ ОТРАСЛЯХ

СУДОМЕТРИКА-2018

МАТЕРИАЛЫ КОНФЕРЕНЦИИ

Санкт-Петербург
15–17 октября 2018 г.

УДК 06.055.2

Материалы Седьмой Международной научно-технической конференции «Измерения и испытания в судостроении и смежных отраслях (СУДОМЕТРИКА-2018)» / ред. В. А. Грановский. – СПб: АО «Концерн «ЦНИИ «Электроприбор», 2018.–190 с.

В сборнике приведены материалы докладов и сообщений, посвященные методам и средствам измерений, испытаний и метрологического обеспечения, а также теории измерений, применительно к сфере создания изделий судостроения.

Под общей редакцией д. т. н., профессора, члена Академии навигации и управления движением, члена Метрологической академии, Заслуженного метролога РФ В. А. Грановского.

СОДЕРЖАНИЕ

УДК 001.8

Статистический анализ результатов измерений: методы, критерии проверки гипотез, проблемы применения и их решение

Б. Ю. Лемешко, П. Ю. Блинов, И. В. Веретельникова, С. Б. Лемешко, А. Ю. Новикова
Кафедра теоретической и прикладной информатики
Новосибирский государственный технический университет
Новосибирск, Россия
Lemeshko@ami.nstu.ru

***Аннотация* — В докладе рассматриваются методы и критерии проверки статистических гипотез, применяемых при статистическом анализе результатов (ошибок) измерений или сличении результатов измерений, выполняемых различными лабораториями. В частности, рассматриваются проблемы применения критериев согласия при проверке сложных гипотез и пути их решения. Приводятся результаты сравнительного анализа множества критериев, используемых при проверке нормальности с указанием достоинств и недостатков рассмотренных критериев, а также совокупности критериев, применяемых для проверки равномерности. Представляются результаты сравнительного анализа совокупности критериев проверки гипотез об однородности законов распределения, предлагаются новые критерии и представляются построенные модели предельных распределений статистик этих критериев. Приводятся результаты сравнительного анализа множества критериев, применяемых при проверке гипотез об однородности средних, и множества параметрических и непараметрических критериев, применяемых при проверке гипотез о равенстве дисперсий. Рассматриваются вопросы разработки программного обеспечения, гарантирующего корректность статистических выводов при использовании рассмотренных множеств критериев.**

Ключевые слова—оценивание параметров; проверка статистических гипотез; критерии согласия; критерии нормальности; критерии равномерности; критерии однородности; мощность критериев

I. Введение

В настоящее время широкий спектр методов статистических анализа в силу ряда объективных и субъективных причин весьма ограничено используется в задачах метрологического обеспечения и анализа результатов измерений. С одной стороны, это связано с тем, что специалисты, занимающиеся измерениями и их анализом, зачастую не знакомы с новыми результатами в силу отсутствия информации, либо с тем, что применение таких методов требует использования сложного программного обеспечения. С другой стороны, причиной такого состояния является и негативный опыт использования статистических методов в условиях нарушения стандартных предположений, обусловливающих корректность применения этих методов. Не последнюю роль играет и то, что реальные свойства критериев проверки статистических гипотез, применяемых при ограниченных объемах выборок, порой существенно отличаются от асимптотических свойств этих критериев.

В данной работе очень кратко остановимся на методах и критериях проверки статистических гипотез, применение которых наиболее перспективно в задачах метрологического обеспечения и анализа результатов измерений, на проблемах, связанных с использованием критериев, и способах их разрешения с опорой на программное обеспечение.

Результаты, о которых будет говориться далее, получены с использованием подхода, предложенного в [1], и развиваемой программной системы ISW [2].

II. Построение моделей

Наиболее часто решаемой задачей является построение параметрической модели закона распределения наблюдаемой случайной величины (ошибки измерения). Классический подход к идентификации закона распределения заключается в последовательной реализации двухэтапной процедуры для каждого вида параметрической модели из рассматриваемого множества законов. На первом этапе по выборочным данным строится модель закона определенного вида, оцениваются параметры этой модели. На втором этапе оценивается степень адекватности полученной модели экспериментальным наблюдениям, как правило, с применением различных критериев согласия.

В качестве оценок параметров могут использоваться оценки максимального правдоподобия (ОМП), MD-оценки, получаемые при минимизации различных расстояний между эмпирической и теоретической функциями распределения, а также оценки, получаемые другими методами. В MD-оценках в качестве расстояний могут использоваться статистики непараметрических критериев согласия (Колмогорова, Крамера–Мизеса–Смирнова, Андерсона–Дарлинга). Наилучшими статистическими свойствами обладают ОМП, но они не являются робастными. Для получения робастных оценок ОМП могут вычисляться по группированным данным.

В качестве моделей законов распределения для наблюдаемых случайных величин (ошибок измерений) в ISW [2] даётся возможность использовать более 3-х десятков параметрических законов, наиболее часто применяемых в приложениях, а также строить модели в виде смеси законов. Описание основной части включенных в [2] параметрических законов приведено в [1].

Заметим, что вид используемых оценок, как правило, отражается на распределениях статистик критериев, используемых при проверке адекватности построенной модели.

Для проверки адекватности могут использоваться различные непараметрические критерии согласия и критерии типа χ^2.

III. Непараметрические критерии согласия

При проверке согласия различают простые и сложные гипотезы. Простая проверяемая гипотеза имеет вид H_0: $F(x) = F(x,\theta)$, где $F(x,\theta)$ – функция распределения вероятностей, с которой проверяют согласие наблюдаемой выборки, а θ – известное значение параметра (скалярного или векторного). Сложная проверяемая гипотеза имеет вид $H_0:\ F(x) \in \{F(x,\theta),\ \theta \in \Theta\}$, где Θ – область определения параметра θ.

Проблемы возникают, если при проверке сложной гипотезы оценку $\hat{\theta}$ параметра распределения вычисляют по той же самой выборке, по которой проверяют согласие. Далее, говоря о проверке сложных гипотез, мы, как правило, будем предполагать, что оценка параметра $\hat{\theta}$ вычисляется по той же выборке.

A. Проверка простых гипотез

Все классические непараметрические критерии согласия предназначены для проверки простых гипотез. При проверке простых гипотез они являются свободными от распределения, то есть распределения статистик критериев при справедливости H_0 не зависят от вида закона $F(x,\theta)$, с которым проверяется согласие.

В **критерии Колмогорова** рекомендуется использовать статистику с поправкой Большева в форме [3]

$$S_K = \sqrt{n}D_n + \frac{1}{6\sqrt{n}} = \frac{6nD_n + 1}{6\sqrt{n}},$$

где $D_n = \max\left(D_n^+, D_n^-\right)$, $D_n^+ = \max\limits_{1\le i\le n}\left\{\frac{i}{n} - F(x_i,\theta)\right\}$;

$D_n^- = \max\limits_{1\le i\le n}\left\{F(x_i,\theta) - \frac{i-1}{n}\right\}$, $x_1, x_2, \ldots, x_n$ – упорядоченные по возрастанию выборочные значения, n – объем выборки. В случае справедливости простой гипотезы распределение статистики сходится равномерно к функции распределения Колмогорова $K(s)$ [3].

Статистика **критерия Крамера–Мизеса–Смирнова** имеет вид

$$S_\omega = \frac{1}{12n} + \sum_{i=1}^{n}\left\{F(x_i,\theta) - \frac{2i-1}{2n}\right\}^2$$

и при справедливости простой гипотезы в пределе подчиняется закону с функцией распределения $a1(s)$ [3].

Статистика **критерия Андерсона–Дарлинга** имеет вид

$$S_\Omega = -n - 2\sum_{i=1}^{n}\left\{\frac{2i-1}{2n}\ln F(x_i,\theta) + \left(1 - \frac{2i-1}{2n}\right)\ln(1 - F(x_i,\theta))\right\}$$

и при справедливости простой проверяемой гипотезы подчиняется закону с функцией распределения $a2(s)$ [3].

В **критерии Купера** целесообразно использовать либо модификацию [4] статистики вида

$$V = (D_n^+ + D_n^-)\left(\sqrt{n} + 0.155 + \frac{0.24}{\sqrt{n}}\right),$$

либо модификацию [5] вида

$$V_n^{mod} = \sqrt{n}(D_n^+ + D_n^-) + \frac{1}{3\sqrt{n}}.$$

Статистика **критерия Ватсона** и используется в форме

$$U_n^2 = \sum_{i=1}^{n}\left(F(x_i,\theta) - \frac{i-1/2}{n}\right)^2 - n\left(\frac{1}{n}\sum_{i=1}^{n}F(x_i,\theta) - \frac{1}{2}\right)^2 + \frac{1}{12n}.$$

Авторами критериев получены распределения статистик при справедливости простой проверяемой гипотезы, вид которых приведен также в [5].

Статистики непараметрических критериев согласия Жанга [6] имеют вид:

$$Z_K = \max_{1\le i\le n}\left(\left(i - \frac{1}{2}\right)\log\left\{\frac{i-1/2}{nF(x_i,\theta)}\right\} + \left(n - i + \frac{1}{2}\right)\log\left[\frac{n-i+1/2}{n\{1-F(x_i,\theta)\}}\right]\right),$$

$$Z_A = -\sum_{i=1}^{n}\left[\frac{\log\{F(x_i,\theta)\}}{n-i+1/2} + \frac{\log\{1-F(x_i,\theta)\}}{i-1/2}\right],$$

$$Z_C = \sum_{i=1}^{n}\left[\log\left\{\frac{[F(x_i,\theta)]^{-1} - 1}{(n-1/2)/(i-3/4) - 1}\right\}\right]^2.$$

Использование этих критериев осложняет сильная зависимость распределений статистик от объема выборки n. Обойти этот недостаток можно за счет использования интерактивного режима исследования распределений статистик применяемых критериев [5].

B. Проверка сложных гипотез

При проверке сложных гипотез свобода от распределения исчезает [7].

При проверке сложных гипотез, когда по той же самой выборке оценивают параметры наблюдаемого закона распределения вероятностей, все рассматриваемые непараметрические критерии согласия теряют свойство «свободы от распределения». Более того, предельные распределения статистик непараметрических критериев согласия зависят от целого ряда факторов, определяющих «сложность» гипотезы.

На закон распределения статистики $G(S|H_0)$ влияют следующие факторы [5]:

– вид наблюдаемого закона распределения $F(x, \theta)$, соответствующего истинной гипотезе H_0;
– тип оцениваемого параметра и число оцениваемых параметров;
– используемый метод оценивания параметров;
– в некоторых ситуациях конкретное значение параметра (например, в случае бета- и гамма-распределений).

Игнорирование того, что проверяют сложную гипотезу, и того, что сложные гипотезы могут быть различными, приводит к некорректному применению непараметрических критериев согласия и, как следствие, к неверным статистическим выводам. Различия в предельных распределениях тех же самых статистик при проверке простых и сложных гипотез настолько существенны, что пренебрегать этим абсолютно недопустимо [5]. Большинство ошибок, совершаемых при использовании непараметрических критериев согласия, связано с незнанием этих обстоятельств. Как правило, при проведении статистического анализа после оценивания по выборке параметров вычисляют значение статистики применяемого критерия, сравнивая его с критическим значением, имеющим место при проверке простой гипотезы, ошибочно не отклоняя проверяемую гипотезу.

За более чем 50 лет с использованием аналитических методов решение проблемы проверки сложных гипотез с использованием непараметрических критериев согласия практически не сдвинулось с места. А вот с использованием компьютерных технологий и методов статистического моделирования [1] удалось построить модели предельных распределений статистик для случая проверки сложных гипотез относительно ряда законов распределений, в основном, для ситуации с использованием ОМП. Построенные модели представлены в руководстве [5]. В ситуациях, когда распределения статистик зависят от конкретных значений оценок параметра или параметров, заранее построить модель предельного распределения статистики критерия невозможно. И здесь выход также кроется в использовании компьютерных технологий, позволяющих в интерактивном режиме (после вычисления оценок параметров и значения статистики критерия) моделировать распределение статистики, а уже по нему находить достигнутый уровень значимости p_{value} и принимать решение о результатах проверки [5]. Именно таким образом решается данная проблема в [2] при использовании непараметрических критериев согласия, в том числе, критериев Жанга. В [5] приведены также результаты сравнительного анализа мощности непараметрических критериев согласия.

IV. Критерии согласия типа χ^2

Статистику критерия согласия χ^2 Пирсона вычисляют по формуле

$$X_n^2 = n\sum_{i=1}^{k}\frac{(n_i / n - P_i(\theta))^2}{P_i(\theta)},$$

где k – число интервалов, n_i – количество наблюдений в интервале, $P_i(\theta)$ – вероятность попадания в интервал.

В случае справедливости простой гипотезы в пределе эта статистика подчиняется χ_r^2-распределению с $r = k-1$ степенями свободы.

В случае проверки сложной гипотезы и справедливости H_0 при условии, что оценки параметров находятся в результате минимизации статистики X_n^2 по этой же самой выборке, статистика X_n^2 асимптотически распределена как χ_r^2 с числом степеней свободы $r = k-m-1$, где m – число оцененных параметров. Это же верно, если использовать ОМП или другие асимптотически эффективные оценки по группированным данным.

При использовании любых оценок по негруппированным данным распределение $G(X_n^2|H_0)$ статистики X_n^2 будет отличаться от χ_r^2-распределения.

Естественно, что ОМП по негруппированным данным более предпочтительны. В этой связи появились критерии Никулина–Рао–Робсона [8, 9], предельными распределениями статистик которых являются χ_{k-1}^2-распределения, а затем и другие модификации критерия. Статистики этих критериев отличаются от статистики критерия X_n^2 только в случае сложной гипотезы и, как правило, несколько выигрывают в мощности. В то же время следует заметить, что использование этих критериев без соответствующего программного обеспечения затруднительно.

Необходимо отметить, что мощность критериев типа χ^2 зависит от числа интервалов и того, как область определения случайной величины разбивается на интервалы. Используя асимптотически оптимальное группирование, можно максимизировать мощность критерия χ^2 Пирсона относительно близких конкурирующих гипотез [1]. В этом случае, при проверке простых гипотез критерий χ^2 Пирсона имеет преимущество в мощности по сравнению с непараметрическими критериями согласия.

V. Критерии проверки нормальности

Принадлежность наблюдаемых данных нормальному закону является необходимой предпосылкой для корректного применения большинства классических методов математической статистики, используемых в задачах обработки измерений, стандартизации и контроля качества. Поэтому проверка на отклонение от нормального закона является частой процедурой в ходе проведения измерений, контроля и испытаний, имеющей особое значение, так как далеко не всегда ошибки измерений, связанные с приборами, построенными на различных физических принципах, или ошибки наблюдений некоторого контролируемого показателя подчиняются нормальному закону. В таких случаях применение классического аппарата, опирающегося на предположение о нормальности наблюдаемого закона, оказывается некорректным и может приводить к неверным выводам.

Действующий ГОСТ [10] охватывает лишь малую часть из существующего множества критериев, ориентированных непосредственно на проверку нормальности, или критериев, которые могут быть использованы для проверки отклонения наблюдаемых данных (ошибок измерений) от нормального закона. В нём не говорится ни об использовании в данных целях критериев типа χ^2, ни об использовании непараметрических критериев согласия. Содержание стандарта и ограниченность перечня включенных в него критериев не позволяет его пользователям ориентироваться в том, какой из критериев наиболее предпочтителен для применения.

В руководстве [11] представлены результаты исследований свойств 15 специальных критериев, ориентированных только на проверку принадлежности выборок нормальному закону, а в системе [2] возможно применение более 25 подобных критериев, не считая рассмотренных выше критериев согласия.

Существенным недостатком значительной части специальных критериев является зависимость распределений статистик от объёмов выборок. Имеющаяся информация о распределениях статистик представлена лишь ограниченными таблицами критических значений. Поэтому по результатам применения критерия нельзя оценить достигнутый уровень значимости p_{value}, а при объёмах выборок больших, чем фигурируют в таблице критических значений, применение критериев оказывается вовсе проблематичным.

Проведенные исследования показали и другие недостатки некоторых критериев. Оказалось, что ряд критериев (например, Шапиро–Уилка, Эппса–Палли, Хегази–Грина и др.) при малых объёмах выборок не способен отличить от нормального конкурирующие законы с чуть более “лёгкими” хвостами. Аналогичную смещённость продемонстрировали и критерии согласия Жанга со статистиками Z_A и Z_K. Сравнительный анализ критериев показал, что некоторые специальные критерии, продемонстрировавшие неудовлетворительные свойства, применять нецелесообразно. И в то же время, при проверке нормальности не следует опираться на какой-то один из критериев, так как нет критерия, обладающего большей мощностью относительно любой альтернативы. В целом, лучшие из специальных критериев показали некоторое преимущество в мощности по сравнению с непараметрическими критериями согласия.

В программной системе [2] при использовании критериев нормальности, вид распределений статистик которых при справедливости H_0 не известен, реализован интерактивный режим моделирования распределений этих статистик при величине n, соответствующей объёму анализируемой выборки, с использованием построенного распределения для оценки p_{value}.

VI. Критерии проверки равномерности

В прикладной математической статистике и в метрологии равномерный закон распределения вероятностей занимает видное место. Именно этим определяется наличие множества специальных критериев, целенаправленно построенных для проверки гипотезы о принадлежности анализируемой выборки равномерному закону (Шермана, Кимбелла, Морана, Ченга–Спиринга, Хегази–Грина, Янга, Фросини, Гринвуда, Гринвуда–Кэсеберри–Миллера, Неймана–Бартона и др.). Для проверки гипотезы о равномерности может использоваться совокупность непараметрических критериев согласия и критерий χ^2 Пирсона. Хотя история применения многих критериев насчитывает уже несколько десятилетий, появляются публикации с предложениями новых критериев равномерности.

В то же время, не смотря на множество публикаций не хватает объективной информации о действительных свойствах критериев, их достоинствах и недостатках. Можно натолкнуться на авторитетные мнения о целесообразности применения тех или иных критериев, которые не подкрепляются результатами сравнительного анализа и не всегда подтверждаются при проверке. Специалистов же, сталкивающихся в своей практической деятельности с необходимостью проверки гипотез о принадлежности наблюдений или ошибок измерений равномерному закону, интересует, какие критерии предпочтительнее использовать и почему.

В руководстве [12] сконцентрированы результаты исследований свойств более 20 специальных критериев, ориентированных только на проверку принадлежности выборок равномерному закону, а также сравнительного анализа этого множества с критериями согласия.

Подбор конкурирующих, используемых при исследовании мощности критериев, позволил оценить достоинства и указать на существенные недостатки некоторых популярных критериев равномерности.

При малых объёмах выборок факт смещённости относительно некоторых альтернатив был отмечен для следующих критериев: Неймана–Бартона (со статистиками N_2, N_3, N_4), Хегази-Грина (со статистиками T_1, T_2, T_1^*, T_2^*), Фросини, Гринвуда–Кэсенберри–Миллера, Кимбелла, Морана 1, Гринвуда, Шермана, Кресси 1, Морана 2, Янга, Дудевича–ван дер Мюлена и 2-х модификаций энтропийного критерия. Этот недостаток не зафиксирован лишь для 4-х критериев (Кресси 2, Шварца, Пардо, Ченга-Спиринга), которые в среднем продемонстрировали не очень высокую мощность. Интересно, что в тех же ситуациях смещённость проявили непараметрические критерии согласия Колмогорова, Крамера–Мизеса–Смирнова, Андрсона–Дарлинга и критерии Жанга.

В то же время анализ мощности при проверке равномерности показал, что непараметрические критерии, особенно, Андерсона–Дарлинга и Жанга со статистикой Z_A не уступают лучшим специальным критериям.

В системе [2] при использовании критериев равномерности, распределения статистик которых представлены таблицами критических значений, оценки p_{value} вычисляются по моделируемым в интерактивном режиме распределениям статистик этих критериев.

VII. Критерии однородности законов

С необходимостью решения задач проверки гипотез о принадлежности двух (или более) выборок случайных величин одной и той же генеральной совокупности (проверки однородности) сталкиваются в различных областях. Такая задача естественно возникает при сличении результатов лабораторных измерений, или аттестации средств измерений, когда пытаются убедиться в том, что закон распределения случайных ошибок не претерпел существенных изменений по истечении некоторого интервала времени.

В настоящей работе мы остановимся только на k-выборочных критериях. Задача проверки однородности k

выборок формулируется следующим образом. Пусть x_{ij} j-е наблюдение i-й выборки $j=\overline{1,n_i}$, $i=\overline{1,k}$. Предполагается, что i-й выборке соответствует непрерывная функция распределения $F_i(x)$. Проверяется гипотеза вида $H_0: F_1(x)=F_2(x)=\ldots=F_k(x)$ без указания общего для них закона распределения.

Общий подход к построению k-выборочных критериев однородности законов, являющихся аналогами двухвыборочных критериев Колмогорова–Смирнова и Крамера–Мизеса (Лемана–Розенблатта), рассматривался в работе [13]. При этом подходе статистика критерия представляет собой меру отклонения эмпирических распределений, соответствующих конкретным выборкам, от эмпирического распределения, построенного по совокупности анализируемых выборок. О k-выборочном варианте критерия Колмогорова–Смирнова, построенного по этому принципу, упоминается в [14, 15]. k-выборочный вариант критерия Андерсона–Дарлинга предложен в [16]. Построенные Жангом в [17] критерии однородности являются развитием критериев однородности Смирнова [3], Лемана–Розенблатта [18, 19] и Андерсона–Дарлинга [20] и позволяют анализировать $k \ge 2$ выборок.

Применение k-выборочных критериев на практике сдерживается тем, что, в лучшем случае, известны лишь критические значения статистик для соответствующих k, как в случае критерия Андерсона–Дарлинга [16] или критериев типа Колмогорова–Смирнова [14, 15], а возможность использования критериев Жанга [17] упирается в необходимость для формирования вывода о результатах проверки гипотезы искать распределения статистик критериев (или оценки достигнутого уровня значимости p_{value}) с использованием статистического моделирования. Исключение составляет лишь критерий однородности χ^2, для которого известны асимптотические распределения статистики при справедливости H_0.

A. Критерий Андерсона–Дарлинга

Статистика k-выборочного критерия Андерсона–Дарлинга в [20] находится следующим образом. По k выборкам строится объединённая упорядоченная выборка $X_1 \le X_2 \le \ldots \le X_n$ объёмом $n=\sum_{i=1}^{k} n_i$, по которой вычисляется

$$A_{kn}^2 = \frac{1}{n}\sum_{i=1}^{k}\frac{1}{n_i}\sum_{j=1}^{n-1}\frac{\left(nM_{ij}-jn_i\right)^2}{j(n-j)},$$

где M_{ij} – число элементов в i-й выборке, которые не больше чем X_j. Статистика критерия имеет вид

$$T_{kn} = \left[A_{kn}^2-(k-1)\right]\Big/\sqrt{D[A_{kn}^2]},$$

где дисперсия определяется выражением

$$D\left[A_{kn}^2\right] = \frac{an^3+bn^2+cn+d}{(n-1)(n-2)(n-3)}$$

при

$$a=(4g-6)(k-1)+(10-6g)H,$$
$$b=(2g-4)k^2+8hk+(2g-14h-4)H-8h+4g-6,$$
$$c=(6h+2g-2)k^2+(4h-4g+6)k+(2h-6)H+4h,$$
$$d=(2h+6)k^2-4hk,$$

где

$$H=\sum_{i=1}^{k}\frac{1}{n_i},\quad h=\sum_{i=1}^{n-1}\frac{1}{i},\quad g=\sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}\frac{1}{(n-i)j}.$$

Распределения статистики T_{kn} в [20] представлены лишь критическими значениями. Асимптотические распределения статистики зависят от числа сравниваемых выборок k и не зависят от n_i. В [21, 22] на основании результатов статистического моделирования были построены модели предельных распределений статистики T_{kn} для $k=2\div 11$. Приведенные в [21] модели построены по смоделированным выборкам статистики при числе имитационных экспериментов $N=10^6$ и $n_i=10^3$.

Критерии однородности Жанга [17] позволяют сравнивать $k \ge 2$ выборок.

Пусть $x_{i1}, x_{i2}, \ldots, x_{in_i}$ упорядоченные выборки непрерывных случайных величин с функциями распределения $F_i(x)$, ($i=\overline{1,k}$) и, как и ранее, $X_1<X_2<\ldots<X_n$, где $n=\sum_{i=1}^{k}n_i$, объединённая упорядоченная выборка. Обозначим R_{ij} ранг j-го упорядоченного наблюдения x_{ij} i-й выборки в объединённой выборке. Пусть $X_0=-\infty$, $X_{n+1}=+\infty$, а ранги $R_{i,0}=1$, $R_{i,n_i+1}=n+1$.

В критериях используется модификация эмпирической функции распределения $\hat{F}(t)$, принимающая в точках разрыва X_m, $m=\overline{1,n}$, значения $\hat{F}(X_m)=(m-0.5)/n$ [17].

Статистика Z_{K} критерия Жанга имеет вид:

B. Критерии Жанга

$$Z_{\text{K}} = \max_{1\le m\le n}\left\{\sum_{i=1}^{k}n_i\left[F_{i,m}\ln\frac{F_{i,m}}{F_m}+\left(1-F_{i,m}\right)\ln\frac{1-F_{i,m}}{1-F_m}\right]\right\},$$

где $F_m=\hat{F}(X_m)$, так что $F_m=(m-0.5)/n$, а вычисление $F_{i,m}=\hat{F}_i(X_m)$ осуществляется следующим образом. В начальный момент значения $j_i=0$, $i=\overline{1,k}$. Если $R_{i,j_i+1}=m$, то $j_i:=j_i+1$ и $F_{i,m}=(j_i-0.5)/n_i$, в противном случае если $R_{i,j_i}<m<R_{i,j_i+1}$, то $F_{i,m}=j_i/n_i$.

Критерий *правосторонний*: проверяемая гипотеза H_0 отклоняется при *больших* значениях статистики.

Статистика Z_{A} критерия задается выражением:

$$Z_{\text{A}} = -\sum_{m=1}^{n}\sum_{i=1}^{k} n_i \frac{F_{i,m}\ln F_{i,m} + \left(1-F_{i,m}\right)\ln\left(1-F_{i,m}\right)}{(m-0.5)(n-m+0.5)},$$

где F_m и $F_{i,m}$ вычисляются, как определено выше.

Критерий *левосторонний*: проверяемая гипотеза отклоняется при *малых* значениях статистики.

Статистика Z_C **критерия Жанга** определяется выражением:

$$Z_C = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\ln\left(\frac{n_i}{j-0.5}-1\right)\ln\left(\frac{n}{R_{i,j}-0.5}-1\right).$$

Этот критерий также *левосторонний*: проверяемая гипотеза H_0 отклоняется при *малых* значениях статистики. Распределения статистик критериев Жанга зависят от объёмов выборок и от числа анализируемых выборок, что осложняет использование критериев.

В то же время, отсутствие информации о законах распределения статистик в современных условиях не является серьёзным недостатком критериев. В программном обеспечении, осуществляющем поддержку применения критериев, несложно организовать вычисление p_{value}, используя методы статистического моделирования, как это и реализуется в [2].

C. Критерии на базе 2-выборочных

Для анализа k выборок можно к каждой паре применить двухвыборочный критерий со статистикой S (всего $(k-1)k/2$ пар), а решение о принятии или отклонении гипотезы H_0 принимать по совокупности результатов. В качестве статистики такого k-выборочного критерия (в случае использования правостороннего двухвыборочного критерия) можно рассмотреть, например, статистику вида

$$S_{\max} = \max_{\substack{1\le i\le k\\ i<j\le k}} \{S_{i,j}\},$$

где $S_{i,j}$ – значения статистик используемого двухвыборочного критерия, вычисленные при анализе i-й и j-й выборок.

Проверяемая гипотеза H_0 будет отклоняться при **больших** значениях статистики $S_{\max}$. Преимуществом такого рода критерия является и то, что в результате будет определена пара выборок, различие между которыми окажется наиболее значимым с позиций используемого двухвыборочного критерия.

В качестве $S_{i,j}$ можно использовать статистики двухвыборочных критериев Смирнова, Лемана–Розенблата, Андерсона–Дарлинга. В этом случае распределения соответствующих статистик $S_{\max}$ сходятся к некоторым предельным (при данном k), модели которых могут быть найдены по результатам статистического моделирования.

Критерий максимума Смирнова. Используемая в критерии Смирнова статистика D_{n_2,n_1} вычисляется в соответствии с соотношениями [3]:

$$D^{+}_{n_2,n_1} = \max_{1\le r\le n_2}\left[\frac{r}{n_2} - F_{1n_1}(x_{2r})\right] = \max_{1\le s\le n_1}\left[F_{2n_2}(x_{2s}) - \frac{s-1}{n_1}\right],$$

$$D^{-}_{n_2,n_1} = \max_{1\le r\le n_2}\left[F_{1n_1}(x_{2r}) - \frac{r-1}{n_2}\right] = \max_{1\le s\le n_1}\left[\frac{s}{n_1} - F_{2n_2}(x_{1s})\right],$$

$$D_{n_2,n_1} = \max\left(D^{+}_{n_2,n_1}, D^{-}_{n_2,n_1}\right).$$

При справедливости гипотезы H_0 и неограниченном увеличении объемов выборок статистика

$$S_{\text{C}} = \sqrt{\frac{n_1 n_2}{n_1+n_2}} D_{n_2,n_1}$$

подчиняется распределению Колмогорова $K(S)$.

В k-выборочном критерии максимума Смирнова в качестве $S_{i,j}$ предпочтительней использовать модификацию статистики Смирнова

$$S_{\text{mod}} = \sqrt{\frac{n_1 n_2}{n_1+n_2}}\left(D_{n_2,n_1} + \frac{n_1+n_2}{4.6 n_1 n_2}\right),$$

распределение которой всегда ближе к предельному распределению Колмогорова $K(S)$ [23]. Статистику $S_{\max}$ в этом случае будем обозначать как $S^{Sm}_{\max}$.

При равных объёмах сравниваемых выборок распределения статистики $S^{Sm}_{\max}$ (как и в двухвыборочном варианте) обладают существенной дискретностью и отличаются от асимптотических (предельных) распределений. Если есть такая возможность, то предпочтительней в качестве n_i выбирать взаимно простые числа, тогда распределения $G(S|H_0)$ статистики $S^{Sm}_{\max}$ практически не отличаются от асимптотических.

Модели асимптотических распределений статистики $S^{Sm}_{\max}$ при числе сравниваемых выборок $k = 3 \div 11$ построены в [24] по результатам статистического моделирования. Хорошими моделями оказались законы семейства бета-распределений 3-го рода с плотностью

$$f(x) = \frac{\theta_2^{\theta_0}}{\theta_3 \mathrm{B}(\theta_0,\theta_1)}\left(\frac{x-\theta_4}{\theta_3}\right)^{\theta_0-1}\left(1-\frac{x-\theta_4}{\theta_3}\right)^{\theta_1-1} \Bigg/ \left[1+(\theta_2-1)\frac{x-\theta_4}{\theta_3}\right]^{\theta_0+\theta_1},$$

представленные в таблице 1 в виде $\mathrm{B_{III}}(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$ при конкретных значениях параметров.

ТАБЛИЦА 1. МОДЕЛИ ПРЕДЕЛЬНЫХ РАСПРЕДЕЛЕНИЙ СТАТИСТИКИ $S_{\max}^{Sm}$

k	**Модель**
2	$K(S)$
3	B_{III} (6.3274, 6.6162, 2.8238, 2.4073, 0.4100)
4	B_{III} (7.2729, 7.2061, 2.6170, 2.3775, 0.4740)
5	B_{III} (7.1318, 7.3365, 2.4813, 2.3353, 0.5630)
6	B_{III} (7.0755, 8.0449, 2.3163, 2.3818, 0.6320)
7	B_{III} (7.7347, 8.6845, 2.3492, 2.4479, 0.6675)
8	B_{III} (7.8162, 8.9073, 2.2688, 2.4161, 0.7120)
9	B_{III} (7.8436, 8.8805, 2.1696, 2.3309, 0.7500)
10	B_{III} (7.8756, 8.9051, 2.1977, 2.3280, 0.7900)
11	B_{III} (7.9122, 9.0411, 2.1173, 2.2860, 0.8200)

В k-выборочном **критерии максимума Лемана–Розенблатта** в качестве $S_{i,j}$ используется статистика соответствующего двухвыборочного критерия, предложенного в работе [18], в форме [3]

$$T = \frac{1}{(n_1+n_2)}\left[n_2\sum_{i=1}^{n_2}\left(r_i - i\right)^2 + n_1\sum_{j=1}^{n_1}\left(s_j - j\right)^2\right] - \frac{4n_1n_2-1}{6(n_1+n_2)},$$

где r_i – порядковый номер (ранг) x_{2i}; s_j – порядковый номер (ранг) x_{1j} в объединенном вариационном ряде. В [19] было показано, что статистика T в пределе распределена как $a1(t)$ [3].

Соответствующую статистику критерия максимума Лемана–Розенблатта обозначим как $S_{\max}^{LR}$.

Построенные в [24] модели асимптотических (предельных) распределений статистики $S_{\max}^{LR}$ при числе сравниваемых выборок $k = 3 \div 11$ представлены в таблице 2. В данном случае наилучшими моделями оказались распределения Sb–Джонсона с плотностью

$$f(x) = \frac{\theta_1\theta_2}{\sqrt{2\pi}(x-\theta_3)(\theta_2+\theta_3-x)}\exp\left\{-\frac{1}{2}\left[\theta_0 - \theta_1\ln\frac{x-\theta_3}{\theta_2+\theta_3-x}\right]^2\right\}$$

при конкретных значениях параметров этого закона, обозначенного в таблице 3 как $Sb(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$. Представленные модели позволяют по значениям статистики $S_{\max}^{LR}$ при соответствующем числе k сравниваемых выборок находить оценки p_{value}.

ТАБЛИЦА 2. МОДЕЛИ ПРЕДЕЛЬНЫХ РАСПРЕДЕЛЕНИЙ СТАТИСТИКИ $S_{\max}^{LR}$

k	**Модель**
2	$a1(t)$
3	Sb(3.2854, 1.2036, 3.0000, 0.0215)
4	Sb(2.5801, 1.2167, 2.2367, 0.0356)
5	Sb(3.1719, 1.4134, 3.1500, 0.0320)
6	Sb(2.9979, 1.4768, 2.9850, 0.0380)
7	Sb(3.2030, 1.5526, 3.4050, 0.0450)

ПРОДОЛЖЕНИЕ ТАБЛИЦЫ 2. МОДЕЛИ ПРЕДЕЛЬНЫХ РАСПРЕДЕЛЕНИЙ СТАТИСТИКИ $S_{\max}^{LR}$

k	Модель
8	Sb(3.2671, 1.6302, 3.5522, 0.0470)
9	Sb(3.4548, 1.7127, 3.8800, 0.0490)
10	Sb(3.4887, 1.7729, 3.9680, 0.0510)
11	Sb(3.4627, 1.8168, 3.9680, 0.0544)

В **критерии максимума Андерсона–Дарлинга** в качестве $S_{i,j}$ используется статистика двухвыборочного критерия Андерсона–Дарлинга, предложенная в [20] и определяемая выражением

$$A^2 = \frac{1}{n_1n_2}\sum_{i=1}^{n_1+n_2-1}\frac{\left(M_i(n_1+n_2)-n_1i\right)^2}{i(n_1+n_2-i)},$$

где M_i – число элементов первой выборки, меньших или равных i-му элементу вариационного ряда объединенной выборки. Предельным распределением статистики A^2 при справедливости проверяемой гипотезы H_0 является распределение $a2(t)$ [3].

Соответствующую статистику k-выборочного критерия максимума Андерсона–Дарлинга обозначим как $S_{\max}^{AD}$.

Построенные в [24] модели асимптотических (предельных) распределений статистики $S_{\max}^{AD}$ при числе сравниваемых выборок $k = 3 \div 11$ представлены в таблице 3. В этом случае лучшими моделями также оказались бета-распределения 3-го рода $\mathrm{B_{III}}(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$ с конкретными значениями параметров.

ТАБЛИЦА 3. МОДЕЛИ ПРЕДЕЛЬНЫХ РАСПРЕДЕЛЕНИЙ СТАТИСТИКИ $S_{\max}^{AD}$

k	**Модель**
2	$a2(t)$
3	B_{III} (4.4325, 2.7425, 12.1134, 8.500, 0.1850)
4	B_{III} (5.2036, 3.2160, 10.7792, 10.000, 0.2320)
5	B_{III} (5.7527, 3.3017, 9.7365, 10.000, 0.3000)
6	B_{III} (5.5739, 3.4939, 7.7710, 10.000, 0.3750)
7	B_{III} (6.4892, 3.6656, 8.0529, 10.500, 0.3920)
8	B_{III} (6.3877, 3.8143, 7.3602, 10.800, 0.4800)
9	B_{III} (6.7910, 3.9858, 7.1280, 11.100, 0.5150)
10	B_{III} (6.7533, 4.2779, 6.6457, 11.700, 0.5800)
11	B_{III} (7.1745, 4.3469, 6.6161, 11.800, 0.6100)

D. Критерий однородности χ^2

Для анализа $k \ge 2$ выборок с успехом может использоваться критерий однородности χ^2. В этом случае общая область, которой принадлежат выборки, разбивается на r интервалов (групп). Пусть η_{ij} – количество элементов i-й выборки, попавших в j-й интервал, тогда $n_i = \sum_{j=1}^{r}\eta_{ij}$.

Статистика критерия однородности χ^2 имеет вид

$$\chi^2 = n\sum_{i=1}^{k}\sum_{j=1}^{r}\frac{(\eta_{ij}-\nu_j n_i/n)^2}{\nu_j n_i} = n\left(\sum_{i=1}^{k}\sum_{j=1}^{r}\frac{\eta_{ij}^{\ 2}}{\nu_j n_i}-1\right),$$

где $\nu_j = \sum_{l=1}^{k}\eta_{lj}$ – общее число элементов всех выборок, попавших j-й интервал. Асимптотическим распределением статистики является χ^2-распределение с числом степеней свободы $(k-1)(r-1)$ [25].

Недостатком критерия является зависимость мощности от выбора числа интервалов и способа разбиения на интервалы.

E. Сравнительный анализ мощности критериев

Мощность k-выборочных критериев исследовалась при разных k относительно различных конкурирующих гипотез.

Кратко результаты исследований можно охарактеризовать следующим образом.

Относительно конкурирующих гипотез, связанных с изменением параметра сдвига у закона распределения одной из выборок, критерии можно упорядочить по мощности следующим образом:

$$S_{\max}^{AD} \succ \mathrm{AD} \succ S_{\max}^{LR} \succ S_{\max}^{Sm} \succ Z_C \succ Z_\mathrm{A} \succ Z_\mathrm{K} \succ \chi^2 .$$

Относительно конкурирующих гипотез, , связанных с изменением параметра масштаба у закона распределения одной из выборок, –

$$Z_C \succ Z_\mathrm{A} \succ Z_\mathrm{K} \succ \mathrm{A\Delta} \succ \chi^2 \succ S_{\max}^{AD} \succ S_{\max}^{Sm} \succ S_{\max}^{LR} .$$

Можно отметить, что с ростом количества сравниваемых выборок тех же объёмов мощность критерия относительно аналогичных конкурирующих гипотез, как правило, снижается.

Отметим также, что критерии Жанга со статистиками Z_K, Z_A, Z_C относительно некоторых альтернатив обладают заметным преимуществом в мощности.

VIII. Критерии однородности средних

К критериям проверки гипотез об однородности математических ожиданий (об однородности средних) прибегают при контроле средств измерений, при статистическом анализе результатов экспериментов, при статистическом управлении качеством для проверки наличия возмущения в ходе процесса.

В общем случае проверяемая гипотеза о равенстве математических ожиданий, соответствующих k выборкам, имеет вид

$$H_0:\ \mu_1 = \mu_2 = \cdots = \mu_k$$

при конкурирующей гипотезе

$$H_1:\ \mu_{i_1} \neq \mu_{i_2}$$

где неравенство выполняется хотя бы для одной пары индексов i_1, i_2.

Для проверки гипотезы H_0 может использоваться ряд параметрических критериев: сравнения двух выборочных средних при известных дисперсиях; сравнения двух выборочных средних при неизвестных, но равных дисперсиях (критерий Стьюдента); сравнения двух выборочных средних при неизвестных и неравных дисперсиях; *F*-критерий. В этих же целях применяется целая совокупность непараметрических критериев: критерий Уилкоксона, критерий Манна–Уитни, критерий Краскела–Уаллиса, критерий Ван дер Вардена, k - выборочный критерий Ван дер Вардена.

Основным предположением, обусловливающим возможность применения параметрических критериев, является принадлежность анализируемых выборок нормальному закону. Непараметрические критерии свободны от этого требования.

Результаты исследований, сконцентрированные в [21], подтвердили устойчивость параметрических критериев проверки однородности математических ожиданий. Если законы распределения анализируемых выборок отличаются от нормального, но нет оснований полагать, что выборки принадлежат законам с «тяжелыми хвостами» или существенно асимметричным, то применение классических параметрических критериев остаётся корректным и не приводит к серьёзным ошибкам при оценке p_{value}.

Многовыборочный критерий Ван дер Вардена практически не уступает в мощности своему параметрическому аналогу *F*-критерию. Недостатком непараметрических критериев Манна–Уитни и Краскела–Уаллиса является существенная дискретность распределений статистик. В то же время оба эти критерия не столь уж сильно уступают в мощности параметрическим критериям.

IX. Критерии однородности средних

Критерии проверки однородности дисперсий очень широко используются в различных приложениях, но неоправданно редко в задачах метрологии.

В критериях проверки однородности дисперсий проверяемая гипотеза о постоянстве дисперсий k выборок имеет вид

$$H_0\!: \sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2 .$$

а конкурирующая с ней –

$$H_1\!: \sigma_{i_1}^2 \neq \sigma_{i_2}^2 ,$$

где неравенство выполняется, по крайней мере, для одной пары индексов i_1, i_2.

В данном случае мы имеем очень представительное множество параметрических критериев (Бартлетта, Кокрена, Хартли, Левене, Фишера, Неймана–Пирсона, О`Брайена, Линка, Ньюмана, Блиса–Кокрена–Тьюки, Кадуэлла–Лесли–Брауна, Оверолла–Вудворда, Миллера, Лайарда) и ряд непараметрических критериев (Ансари–Бредли, Муда, Сижела–Тьюки, Клотца, Кейпена, Флайне–Киллина). Одним из основных предположений при построении параметрических критериев является принадлежность анализируемых выборок нормальному закону распределения. Как правило, параметрические критерии чрезвычайно чувствительны к малейшим нарушениям этого предположения. Именно попытками построения устойчивых к нарушению стандартного предположения о нормальности объясняется большой выбор параметрических критериев. Нельзя не отметить, что повышение устойчивости критерия практически всегда сопровождается некоторой потерей в мощности.

Непараметрические аналоги предназначены для проверки гипотез об однородности параметров масштаба законов. Корректность применения непараметрических критериев обусловлена не менее жесткими предположениями. Во-первых, предполагается, что анализируемые выборки принадлежат одному и тому же закону распределения. В противном случае даже при справедливости проверяемой гипотезы H_0 о равенстве дисперсий распределения статистик этих критериев будут отличаться от имеющих место при выполнении данного предположения. Во-вторых, предполагается равенство математических ожиданий.

Результаты наших исследований свойств критериев проверки гипотез об однородности дисперсий, сконцентрированные в [21], кратко можно охарактеризовать следующим образом.

При $k = 2$ и выполнения предположения о нормальности наиболее предпочтительна по мощности группа эквивалентных критериев Бартлетта, Кокрена, Хартли, Фишера, Неймана–Пирсона и.

Другую группу практически эквивалентных по мощности критериев образуют критерии Лайарда, Миллера, О'Брайена и модифицированный критерий Оверолла–Вудворда. При этом эта группа имеет преимущество в мощности по сравнению с критерием Левене. Считается, что, как и последний, критерии О'Брайена и модифицированный критерий Оверолла–Вудворда достаточно устойчивы к нарушению стандартного предположения о нормальности.

Непараметрические критерии существенно уступают в мощности вышеперечисленным параметрическим, но, как правило, превосходят в мощности параметрические критерии Ньюмана, Блисса–Кокрена–Тьюки, Кадуэлла–Лесли–Брауна и Линка. В группе непараметрических критериев результаты анализа показывают заметное преимущество в мощности критерия Клотца. Затем следуют критерий Флайне-Киллина и критерий Муда. Ещё меньшую мощность и практическую эквивалентность демонстрируют критерии Ансари–Бредли и Сижела–Тьюки.

При $k > 2$ критерии Бартлетта, Кокрена, Хартли, Неймана–Пирсона и критерий Оверолла–Вудворда уже не образуют группу эквивалентных критериев с одинаковой мощностью. Исключение составляют лишь критерии Бартлетта и Неймана–Пирсона, которые остаются практически эквивалентными, а также пара критериев Миллера и Лайарда. На первой позиции с явным преимуществом оказывается критерий Кокрена.

Применению параметрических критериев в нестандартных ситуациях при законах отличных от нормального препятствует то, что распределения статистик критериев при справедливости H_0 оказываются неизвестными. Но это не исключает возможности их применения при использовании компьютерных технологий для исследования распределений статистик в интерактивном режиме.

В случае принадлежности выборок симметричному закону с "хвостами", более легкими, чем у нормального закона, преимущество в мощности сохраняется за параметрическими критериями, и критерии упорядочиваются по практически так же, как и при нормальном законе.

При симметричных законах с более тяжелыми "хвостами" по сравнению с нормальным законом порядок предпочтения меняется, и при $k = 2$ критерии располагаются следующим образом:

Флайне–Киллина $\succ$ *Клотца* $\succ$ *Муда* $\succ$ *Левене* $\succ$ *Сижела–Тьюки* $\sim$ *Ансари–Бредли* $\succ$ *Миллера* $\succ$ *О'Брайена* $\succ$ *Лайарда* $\succ$ *Модифицированный Оверолла–Вудворда* $\succ$ ***группа критериев*** *(Бартлетта, Кокрена, Хартли, Фишера, Неймана–Пирсона, Оверолла–Вудворда)* $\succ$ *Ньюмана* $\succ$ ***группа критериев*** *(Блисса–Кокрена–Тьюки, Кадуэлла–Лесли–Брауна, Линка)*

В той же ситуации, но при $k > 2$, когда могут применяться не все критерии, имеем следующую упорядоченность:

Флайне–Киллина $\succ$ *Левене* $\succ$ *О'Брайена* $\succ$ *Лайарда* $\succ$ *Модифицированный Оверолла–Вудворда* $\succ$ *Миллера* $\succ$ *Бартлетта* $\sim$ *Неймана–Пирсона* $\succ$ *Оверолла–Вудворда* $\succ$ *Хартли* $\succ$ *Кокрена* $\succ$ *Кадуэлла–Лесли–Брауна* $\succ$ *Блисса–Кокрена–Тьюки.*

В системе [2] реализована возможность применения параметрических критериев проверки гипотез об однородности дисперсий при законах распределения, отличающихся от нормального, но известных. Информация о предполагаемом законе может быть следствием предшествующих исследований. Она может быть получена также в результате идентификации закона по выборке, полученной объединением анализируемых.

Более полная информация о достоинствах и недостатках конкретных критериев представлена в [21].

X. Заключение

Очевидно, что в задачах метрологического обеспечения измерительных процессов используется далеко не полный спектр эффективных статистических методов, способных оказать положительное влияние на качество этого обеспечения. Осознание этого факта и знание реальных свойств множества статистических критериев должно способствовать прогрессу в решении задач, стоящих перед метрологами.

Далеко не последняя роль в разрешении множества проблем, препятствующих эффективному использованию статистических критериев, в том числе в задачах метрологии, принадлежит компьютерным технологиям исследований и программному обеспечению.

Благодарности

Исследования выполнены при поддержке Министерства образования и науки РФ в рамках государственной работы «Обеспечение проведения научных исследований» (№ 1.4574.2017/6.7) и проектной части государственного задания (№ 1.1009.2017/4.6).

ЛИТЕРАТУРА

[1] Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н., Чимитова Е.В. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: Монография. Новосибирск : Изд-во НГТУ, 2011. 888 с.

[2] ISW–Программная система статистического анализа одномерных наблюдений. https://ami.nstu.ru/~headrd/ISW.htm. (дата обр. 06.08.2018)

[3] Большев Л.Н. Смирнов Н.В. Таблицы математической статистики. – М. : Наука, 1983. 416 с.

[4] Stephens M. A. EDF statistics for goodness of fit and some comparisons // J. Am. Statist. Assoc. 1974. Vol. 69. No. 347. pp. 730–737.

[5] Лемешко Б.Ю. Непараметрические критерии согласия: Руководство по применению: Монография. – М.: ИНФРА-М, 2014. 163 с.

[6] Zhang J. Powerful goodness-of-fit tests based on the likelihood ratio // Journal of the Royal Statistical Society: Series B. 2002. V. 64. № 2. pp. 281-294.

[7] M. Kac, J. Kiefer, J. Wolfowitz. On tests of normality and other tests of goodness of fit based on distance methods // Ann. Math. Stat. 1955. Vol. 26. pp. 189–211.

[8] Никулин М.С. О критерии хи-квадрат для непрерывных распределений // Теория вероятностей и ее применение. 1973. Т. XVIII. № 3. С. 675–676.

[9] Rao K.C., Robson D.S. A chi-squared statistic for goodness-of-fit tests within the exponential family // Commun. Statist. 1974. Vol. 3. pp. 1139–1153.

[10] ГОСТ Р ИСО 5479–2002. Статистические методы. Проверка отклонения распределения вероятностей от нормального распределения. – М. : Изд-во стандартов, 2002. 30 с.

[11] Лемешко Б.Ю. Критерии проверки отклонения распределения от нормального закона. Руководство по применению: Монография. – М.: ИНФРА-М, 2015. 160 с.

[12] Лемешко Б.Ю., Блинов П.Ю. Критерии проверки отклонения распределения от равномерного закона. Руководство по применению: Монография. – М.: ИНФРА-М, 2015. 183 с.

[13] Kiefer J. K-Sample Analogues of the Kolmogorov-Smirnov and Cramer-v. Mises Tests // Annals of Mathematical Statistics. 1959. Vol. 30. No. 2. pp. 420-447.

[14] Conover W.J. Several k-sample Kolmogorov-Smirnov tests // The Annals of Mathematical Statistics. 1965. Vol. 36, No. 3. pp.1019-1026.

[15] Conover W.J. Practical Nonparametric Statistics. 3d ed. Wiley, 1999. 584 p.

[16] Scholz F.W., Stephens M.A. K-Sample Anderson–Darling Tests // Journal of the American Statistical Association. 1987. Vol. 82. No. 399. pp. 918-924.

[17] Zhang J., Wu Y. k-Sample tests based on the likelihood ratio // Computational Statistics & Data Analysis. 2007. V. 51. No. 9. pp. 4682-4691.

[18] Lehmann E.L. Consistency and unbiasedness of certain nonparametric tests // Ann. Math. Statist. 1951. Vol. 22, № 1. pp. 165–179.

[19] Rosenblatt M. Limit theorems associated with variants of the von Mises statistic // Ann. Math. Statist. 1952. Vol. 23. pp. 617–623.

[20] Pettitt A.N. A two-sample Anderson-Darling rank statistic // Biometrika. 1976. Vol. 63. No. 1. pp. 161-168.

[21] Лемешко Б.Ю. Критерии проверки гипотез об однородности. Руководство по применению : монография. – М. : ИНФРА-М, 2017. 208 с.

[22] Лемешко Б.Ю., Лемешко С.Б., Веретельникова И.В. О применении критериев проверки однородности законов распределения // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2017. № 41. С. 24-31.

[23] Lemeshko B.Yu., Lemeshko S.B. Statistical distribution convergence and homogeneity test power for Smirnov and Lehmann–Rosenblatt tests // Measurement Techniques, 2005. V. 48, № 12. pp. 1159-1166.

[24] Лемешко Б.Ю., Веретельникова И.В. Мощность k-выборочных критериев проверки однородности законов // Измерительная техника. 2018. № 7. С. 3-7.

[25] Крамер Г. Математические методы статистики. – М.: Мир, 1975. 648 с.