Математические методы исследования
Mathematical methods of investigation

DOI: https://doi.org/10.26896/1028-6861-2019-85-5-67-79

МЕТОД СТАТИСТИЧЕСКИХ ИСПЫТАНИЙ В ПРИКЛАДНОЙ СТАТИСТИКЕ

© **Александр Иванович Орлов**

Московский государственный технический университет им. Н. Э. Баумана, Россия, 105005, Москва, Бауманская 2-я, д. 5; e-mail: prof-orlov@mail.ru

Статья поступила 25 декабря 2018 г. Поступила после доработки 17 января 2019 г. Принята к публикации 25 февраля 2019 г.

Новая парадигма математических методов исследования опирается на эффективное применение информационно-коммуникационных технологий как при расчете характеристик методов анализа данных, так и при имитационном моделировании. Датчики псевдослучайных чисел лежат в основе многих современных технологий анализа данных. Для решения конкретных прикладных задач исследователи постоянно разрабатывают новые методы обработки статистических данных — результатов измерений (наблюдений, испытаний, анализов, опытов) и экспертных оценок. Свойства каждого вновь предлагаемого метода необходимо изучить. Интеллектуальными инструментами являются предельные теоремы и метод статистических испытаний (Монте-Карло). В 2016 г. наш журнал начал дискуссию о современном состоянии и перспективах развития статистического моделирования, т.е. теории и практики применения метода статистических испытаний (Монте-Карло), различных вариантов имитационного моделирования. Предыдущая дискуссия о свойствах таких датчиков была проведена в нашем журнале в 1985 – 1993 гг. Данная статья посвящена применению метода статистических испытаний для изучения свойств статистических критериев проверки однородности двух независимых выборок. Рассмотрены: критерий Крамера – Уэлча, совпадающий при равенстве объемов выборок с критерием Стьюдента; критерии Лорда, Вилкоксона (Манна – Уитни), Вольфовица, Ван-дер-Вардена, Смирнова, типа омега-квадрат (Лемана – Розенблатта). В качестве функций распределения элементов двух выборок заданы нормальные распределения и распределения Вейбулла – Гнеденко. Установлено, что для проверки гипотезы совпадения функций распределения двух выборок целесообразно использовать критерий Лемана – Розенблатта типа омега-квадрат. Если есть основания предполагать, что распределения отличаются в основном сдвигом, то можно использовать критерии Вилкоксона и Ван-дер-Вардена. Однако даже в этом случае критерий омега-квадрат может оказаться более мощным. В общем случае, кроме критерия Лемана – Розенблатта, допустимо применение критерия Смирнова — с учетом отличия реального уровня значимости от номинального. Изучены частоты расхождений статистических выводов по разным критериям.

Ключевые слова: прикладная статистика; метод статистических испытаний; метод Монте-Карло; датчики псевдослучайных чисел; критерии проверки статистических гипотез; однородность двух независимых выборок; вычислительный эксперимент; критерий Крамера – Уэлча; критерий Лорда; критерий Вилкоксона; критерий Ван-дер-Вардена; критерий Смирнова; критерий Лемана – Розенблатта.

STATISTICAL SIMULATIONS METHOD IN APPLIED STATISTICS

© **Alexander I. Orlov**

Bauman Moscow State Technical University, 2-ya Baumanskaya ul., 5, 105005, Moscow, Russia, e-mail: prof-orlov@mail.ru

Received December 25, 2018. Revised January 17, 2019. Accepted February 25, 2019.

The new paradigm of mathematical research methods is based on the effective application of information and communication technologies both in calculating the characteristics of the methods of data analysis and in simulation modeling. Pseudo-random number generators underlie many modern data analysis technologies. To solve specific applied problems, researchers permanently develop the new methods for processing statistical data, i.e., measurement results (observations, tests, analyzes, experiments) and expert estimations. The properties of each newly proposed method must be studied. The intellectual tools are limit theorems and method of statistical simulations (Monte-Carlo method). In 2016, our journal opened a discussion on the current state and prospects for the development of statistical modeling, i.e., the theory and practice of applicating the method of the statistical simulations (Monte-Carlo method), and various variants of the simulation. The previous discussion about the properties of such generators was conducted in our journal in 1985 – 1993. This article is devoted to application of the statistical simulations method to the study of the properties of statistical criteria for testing the homogeneity of two independent samples. We consider: the Kramer – Welch criterion, which coincides with Student's criterion when sample sizes are equal; the criteria of Lord, Wilcoxon (Mann – Whitney), Wolfowitz, Van der Waerden, Smirnov, ω^2 (Lehmann – Rosenblatt). It is necessary to set the distribution functions of the elements of two samples. We use the normal and Weibull – Gnedenko distributions. It is shown advisable to use the Lehmann – Rosenblatt ω^2 test when testing the hypothesis of coincidence of the distribution functions of two samples. If there is a reason to assume that the distributions differ mainly in the shift, then the Wilcoxon and Van der Waerden criteria can be used. However, even in this case, the ω^2 test may be more powerful. In the general case, apart from the Lehmann – Rosenblatt criterion, the use of the Smirnov criterion is permissible, taking into account the difference between the real level of significance and the nominal one. The frequency of the discrepancies of statistical findings based on different criteria is studied.

Keywords: applied statistics; statistical simulations method; Monte-Carlo method; pseudo-random number generators; criteria for testing statistical hypotheses; homogeneity of two independent samples; computational experiment; Kramer – Welch criterion; Lord criterion; Wilcoxon criterion; Van der Waerden criterion; Smirnov criterion; Lehmann – Rosenblatt criterion.

Новая парадигма математических методов исследования

В развитии математических методов исследования выделяем два важных периода [1]. Первый — начало XX в., когда были разработаны базовые положения современной математической статистики, сформулированы основные идеи таких ее разделов, как описание данных, оценивание параметров, проверка статистических гипотез. Эти идеи легли в основу учебников, используемых и в настоящее время. Наряду с рациональными приемами анализа данных продолжают пропагандироваться устаревшие воззрения, например, основанные на использовании параметрических семейств распределений вероятностей, в то время как установлено, что практически все распределения реальных данных не являются нормальными и не описываются с помощью иных семейств распределений вероятностей. Второй период — с 1980-х годов по настоящее время. Усилиями сотен исследователей разработана новая парадигма прикладной статистики [2]. Фактически речь идет о новой парадигме математических методов исследования [3]. В соответствии с новой парадигмой заложены основы математики XXI в. — системной нечеткой интервальной математики [4]. На первое место вышла статистика нечисловых данных. Так, за десять лет (2006 – 2015 гг.) ей посвящено 27,6 % всех публикаций раздела «Математические методы исследования» нашего журнала, т.е. 63,0 % среди статей по прикладной статистике [5].

Новая парадигма математических методов исследования опирается на эффективное применение информационно-коммуникационных технологий как при расчете характеристик методов анализа данных, так и при имитационном моделировании. Датчики псевдослучайных чисел лежат в основе многих современных технологий анализа данных. Эти эффективные инструменты исследователя внутренне противоречивы — в них с помощью детерминированных алгоритмов получаем последовательность чисел, обладающих многими свойствами случайных величин. Поэтому свойства таких инструментов требуют тщательного изучения.

Метод статистических испытаний — инструмент исследователя

Для решения конкретных прикладных задач исследователи постоянно разрабатывают новые методы обработки статистических данных — результатов измерений (наблюдений, испытаний, анализов, опытов) и экспертных оценок. Свойства каждого вновь предлагаемого метода необходимо изучить. Какие интеллектуальные инструменты можно применить для такого изучения?

Мощным инструментом исследователей в области математической статистики являются предельные теоремы теории вероятностей — закон больших чисел, центральная предельная теорема и т.п. Некоторые ориентированные на математику специалисты призывают ими и ограничиться. Однако для практического использования статистических методов предельных теорем недостаточно. Необходимо выяснить, начиная с какого объема выборки можно пользоваться результатами, полученными с помощью предельных теорем, и понять, как принимать решения, если объем имеющихся данных меньше этой границы.

С середины XX в. исследователю доступна универсальная «отмычка» — метод статистических испытаний (метод Монте-Карло), другими словами, имитационное моделирование. Он основан на использовании последовательности псевдослучайных чисел, свойства которых напоминают свойства рассматриваемых в теории вероятностей случайных величин. Основная идея состоит в последовательном выполнении следующих этапов: 1) разработка вероятностно-статистической модели реального явления или процесса; 2) планирование статистического испытания, в котором случайные величины заменяются псевдослучайными, полученными с помощью того или иного датчика псевдослучайных чисел; 3) проведение большого числа испытаний (тысяч или миллионов); 4) анализ полученных результатов расчетов.

С каждым этапом связаны соответствующие проблемы адекватности имитационного моделирования. Так, для предельных теорем обычно справедлив тот или иной принцип инвариантности, т.е. в пределе исчезает зависимость от конкретного вида распределения. Однако при изучении скорости сходимости выбор этого конкретного вида весьма важен, поскольку от него зависит итоговый результат статистического моделирования: один — для нормального распределения, другой — для логистического, третий — для распределения Коши...

Датчики псевдослучайных чисел лишь имитируют случайность. Алгоритмы получения псевдослучайных чисел имеют достаточно короткое описание, в то время как по определению А. Н. Колмогорова (в рамках теории информации) описание случайной последовательности должно расти пропорционально длине этой последовательности [6]. Кроме этой глобальной причины методологической несостоятельности датчиков псевдослучайных чисел, есть и частные недостатки. Например, у некоторых популярных до настоящего времени датчиков три последовательных значения связаны линейной зависимостью.

Значения, рассчитанные с помощью метода Монте-Карло, имеют погрешности, определяемые конечностью числа испытаний. При оценивании вероятности события погрешность достигает величины $1/(2\sqrt{N})$, где N — число испытаний. Значит, для оценивания вероятности с точностью 10^{-6} необходимо $10^{12}/4$ испытаний. На практике провести такое количество испытаний невозможно.

Дискуссия о современном состоянии и перспективах развития статистического моделирования

Проблемы теории и практики статистических испытаний (Монте-Карло) заслуживают тщательного обсуждения. В 2016 г. наш журнал начал дискуссию о современном состоянии и перспективах развития статистического моделирования, т.е. теории и практики применения метода статистических испытаний (Монте-Карло), различных вариантов имитационного моделирования. Предыдущая дискуссия о свойствах таких датчиков была проведена в нашем журнале в 1985 – 1993 гг.

«Затравкой» дискуссии послужили статьи [7] и [8]. В первой из них рассмотрены задачи повышения эффективности вычислений методом Монте-Карло. Отмечено, что ключевую роль в их решении играют вопросы выбора объема статистических испытаний (количества моделируемых случайных чисел), а также качества соответствующих датчиков случайных чисел. Обсуждены проблемы реализации алгоритмов методов Монте-Карло, обусловленные требованиями повышения скорости сходимости асимптотических решений к истинным решениям.

В статье [8] констатируется, что цель прикладной математической статистики — разработка методов анализа данных, предназначенных для решения конкретных прикладных задач. С течением времени подходы к разработке таких методов менялись. Сто лет назад принимали, что распределения данных имеют определенный вид, например, являются нормальными, и исходя из этого предположения развивали статистическую теорию. На следующем этапе на первое место в теоретических исследованиях выдвинулись предельные теоремы. Под «малой выборкой» понимают такую выборку, для которой нельзя применять выводы, основанные на предельных теоремах. В каждой конкретной статистической задаче возникает необходимость разделить конечные объемы выборки на два класса: для одного можно применять предельные теоремы, а для другого делать этого нельзя из-за риска получения неверных выводов. Для выбора границы часто используют метод Монте-Карло (статистических испытаний). Более сложные проблемы возникают при изучении влияния на свойства статистических процедур анализа данных тех или иных отклонений от исходных предположений. Такое влияние также часто изучают, используя метод Монте-Карло. Основная и пока не решенная в общем виде проблема при изучении устойчивости выводов при наличии отклонений от параметрических семейств распределений состоит в том, какие распределения использовать для моделирования. Сформулированы и другие нерешенные проблемы [8].

Подборка из трех статей опубликована в третьем номере 2017 г. О. И. Кутузов и Т. М. Татарникова [9] рассмотрели две задачи, обусловленные особенностями применения имитационного моделирования при исследовании сложных технических систем. Одна из них связана с реализацией подхода к повышению эффективности метода Монте-Карло при моделировании редких событий: сочетание расслоенной выборки с равновзвешенным моделированием позволяет значительно ускорить алгоритмический анализ моделей стохастических систем методом имитации. Решение другой задачи выявило проблему, связанную с неадекватностью использования одного и того же датчика псевдослучайных чисел при сопоставлении выборочных значений очередей, полученных на имитационных моделях фрактальной и классической систем массового обслуживания.

И. З. Аронов и О. В. Максимова [10] представили результаты статистического моделирования, характеризующие зависимость времени достижения консенсуса от числа членов технических комитетов по стандартизации (ТК) и их авторитарности. Использована математическая модель обеспечения консенсуса в работе ТК, основанная на модели, предложенной Де Гроотом. Проведен анализ основных проблем достижения консенсуса при разработке консенсусных стандартов в условиях предложенной модели. Показано, что увеличение числа экспертов ТК и их авторитарности негативно влияет на время достижения консенсуса и способствует разобщенности группы.

В комментарии [11] к этой статье проанализировано соотношение консенсуса и истины. Работа технических комитетов по стандартизации — одна из форм экспертных процедур, поэтому ее целесообразно рассматривать в рамках теории и практики экспертных оценок. Тогда проблема консенсуса — это проблема согласованности мнений членов комиссии экспертов. Однако цель работы экспертной комиссии — не достижение согласованности экспертов (консенсуса), а получение (в качестве коллективного мнения) выводов, отражающих реальность, обычно нацеленных на выработку обоснованных управленческих решений, короче говоря, на получение истины. Наблюдаем объективное противоречие между стремлением к выявлению истины и желанием обеспечить консенсус.

Итоги первого этапа дискуссии подведены в [12]. Опубликован ряд статей, посвященных применению метода статистических испытаний (Монте-Карло) для решения различных задач. Так, М. С. Жуков применяет его для изучения свойств алгоритмов нахождения медианы Кемени как итогового мнения комиссии экспертов [13], а И. В. Гадолина и Н. Г. Лисаченко — при разработке метода построения доверительных интервалов для процентилей случайной выборки прочности композитов [14]. Столь интересно начатая дискуссия заслуживает продолжения и расширения круга обсуждаемых проблем.

Статистические критерии проверки однородности двух независимых выборок

Обсудим применение метода статистических испытаний для изучения свойств статистических критериев проверки однородности двух независимых выборок.

Исходные данные — две выборки — $x_1, x_2, \ldots, x_m$ и $y_1, y_2, \ldots, y_n$ (т.е. наборы из m и n действительных чисел), требуется проверить их однородность.

В общепринятой модели $x_1, x_2, \ldots, x_m$ — независимые одинаково распределенные случайные величины с функцией распределения $F(x)$, а $y_1, y_2, \ldots, y_n$ — также независимые одинаково распределенные случайные величины, но, вообще говоря, с другой функцией распределения $G(x)$.

Разделяют однородность характеристик (равенство или математических ожиданий, или медиан, или дисперсий и т.п.) и однородность (совпадение) функций распределения (абсолютную однородность). Во втором случае речь идет о проверке нулевой гипотезы:

$$H_0\colon F(x) = G(x) \text{ при всех } x.$$

Отсутствие однородности означает, что верна альтернативная гипотеза, согласно которой

$$H_1\colon F(x_0) \neq G(x_0) \text{ хотя бы при одном значении аргумента } x_0.$$

Если гипотеза H_0 принята, то выборки можно объединить в одну, если нет, то нельзя.

Рассмотрим следующие статистические критерии, предназначенные для проверки однородности двух независимых выборок.

1. Критерий Крамера – Уэлча T, совпадающий при равенстве объемов выборок ($m = n$) с критерием Стьюдента t [15].

2. Критерий Лорда, или модифицированный t-критерий [16, табл. 3.10, с. 42] со статистикой

$$L = \frac{2|\bar{x} - \bar{y}|}{\left(\max\limits_{1\le i\le m} x_i - \min\limits_{1\le i\le m} x_i\right) + \left(\max\limits_{1\le i\le n} y_i - \min\limits_{1\le i\le n} y_i\right)}.$$

3. Критерий Вилкоксона (Манна – Уитни) ([16, табл. 6.8, с. 94], [17]), основанный на статистике U-сумме рангов элементов первой выборки в общем вариационном ряду.

4. Критерий Вольфовица V, основанный на количестве серий в общем (объединенном) вариационном ряду (серия — часть последовательности, состоящая из элементов одной выборки) и разобранный в [16, табл. 6.7].

5. Критерий Ван-дер-Вардена [16, табл. 6.9], основанный на статистике

$$X = \sum_{i=1}^{m} \Psi\left\{\frac{r_i}{m+n+1}\right\},$$

где r_i — ранг i-го элемента первой выборки в общем вариационном ряду; $\Psi(t)$ — функция, обратная к функции стандартного нормального распределения $\Phi(x)$.

6. Критерий Смирнова [16, 18], основанный на статистике

$$S = D_{m,n} = \sup_x |F_m(x) - G_n(x)|,$$

где $F_m(x)$ и $G_n(x)$ — эмпирические функции распределения, построенные по первой и второй выборкам.

7. Критерий типа омега-квадрат [16, 18], предложенный Леманом [19], изученный впервые Розенблаттом [20], а потому называемый критерием Лемана – Розенблатта. Этот критерий основан на статистике

$$\omega^2 = \omega^2_{mn} = \frac{mn}{m+n} \int_{-\infty}^{+\infty} (F_m(x) - G_n(x))^2 dH_{m+n}(x),$$

где $H_{m+n}(x)$ — эмпирическая функция распределения, построенная по объединенной выборке.

За пределами перечня остались многие критерии — хи-квадрат [21], Сэвиджа [22], знаков [23], основанные на последовательных рангах [24], и др.

Постановка задачи изучения статистических критериев методом статистических испытаний

С помощью вычислительных экспериментов по изучению свойств критериев однородности двух выборок можно выяснить, при каких объемах выборок следует пользоваться предельными распределениями. Ясно, что ответ определяется заданной исследователем точностью (максимально возможным отклонением допредельного распределения от предельного на заданном отрезке или на всей прямой). Можно сравнивать критерии по мощности при тех или иных конкретных альтернативах (например, альтернативах сдвига или масштаба). Представляет интерес анализ «корреляции» критериев на основе изучения доли совпадающих решений по результатам проверки статистических гипотез с помощью этих критериев (эта задача допускает несколько вариантов постановок — можно сравнивать критерии при фиксированном уровне значимости, например 0,05, можно использовать несколько уровней значимости, можно установить связь между достигаемыми уровнями значимости, ...).

Поскольку статистики ранговых критериев принимают лишь конечное число значений, то их распределения дискретны. Поэтому они «проскакивают» обычно используемые в таблицах [16, 23] номинальные уровни значимости — 0,01; 0,05; 0,1 и др. Особенно существенным это обстоятельство оказывается для статистик, принимающих небольшое число значений, таких как статистика Смирнова: реальный уровень значимости статистического критерия может быть в несколько раз меньше номинального, например, равняться 0,02 вместо 0,05 [21, 25]. Сравнение непараметрических критериев затрудняется тем, что по указанной причине невозможно обеспечить совпадение их уровней значимости. Казалось бы, можно использовать рандомизированные критерии. Однако их использование не соответствует большинству практических задач, в которых проверяется однородность двух конкретных выборок. Рандомизированные критерии нацелены на обработку большого числа однотипных выборок фиксированных объемов.

Таким образом, многообразие перспективных вычислительных экспериментов обширно. Для обеспечения изучения свойств различных критериев проверки гипотез однородности нами совместно с Ю. Э. Камнем и Я. Э. Камнем разработан программный продукт, состоящий из четырех блоков: генерации равномерно распределенных на [0; 1] псевдослучайных чисел; вычисления на их основе псевдослучайных чисел с заданными законами распределения; расчета значений статистик критериев; блока сервисных и управляющих программ.

При моделировании использовался датчик равномерно распределенных на множестве $\{1, 2, ..., 2^{15} - 1\}$ псевдослучайных чисел [26], построенный на основе рекуррентной формулы

$$x_{n+1} = (1285x_n + 6925)\mathrm{mod}(2^{15}),\ n = 1, 2, ... \quad (1)$$

Тестирование [27] этого датчика с помощью критерия Колмогорова для выборок объема 5000 на уровне значимости 2,5 % показало согласие с равномерным распределением. Поскольку далее гипотеза однородности проверяется при уровне значимости 0,05, то погрешность метода Монте-Карло оценивается как $\pm\sqrt{\frac{0{,}05 \cdot 0{,}95}{5000}} = \pm\ 0{,}003$.

Академик АН СССР Ю. В. Прохоров при неформальном обсуждении проблем статистического моделирования, проведенном в рамках Первого Всемирного конгресса Общества математической статистики и теории вероятностей им. Бернулли [28], отмечал, что применения метода Монте-Карло можно разделить на два класса. В первом из них, появившемся исторически раньше, качество датчика определяется соответствием распределения даваемых датчиком псевдослучайных чисел заданному распределению, например равномерному. Выполнения этого условия достаточно, в частности, для вычисления многомерных интегралов. Именно этот класс применений обычно имеется в виду в литературе по методу Монте-Карло [29, 30]. Для применений второго класса важно обеспечить независимость псевдослучайных чисел, точнее, достаточное для успешного применения датчика приближение к независимости. Как показано в работах И. Г. Журбенко с соавторами [31 – 33], датчики типа (1) принципиально не могут обеспечить независимость. Однако из расчетов Г. В. Рыдановой [34] следует, что последовательности из не более чем 24 псевдослучайных чисел, используемые для одного статистического испытания, могут рассматриваться как модели последовательностей независимых случайных величин.

Мы активно используем метод Монте-Карло в научных исследованиях. В частности, для изучения скорости сходимости распределений статистик — в предельной теории помех, создаваемых электровозами [35], в теории люсианов [36], при изучении свойств критериев однородности [25]. Но одновременно отдаем себе отчет в недостатках этого инструмента и предостерегаем от его бездумного употребления [37].

Для постановки вычислительного эксперимента необходимо задать функции распределения элементов двух выборок — $F(x)$ и $G(x)$. Обоснованных теорией или практикой рекомендаций по выбору $F(x)$ и $G(x)$ в настоящее время нет. Поэтому для поискового исследования будем использовать привычные нормальные распределения и распределения Вейбулла – Гнеденко.

Функция распределения Вейбулла – Гнеденко имеет вид

$$F(x;a,b,c)=\begin{cases}1-\exp\left[-\left(\dfrac{x-a}{b}\right)^{c}\right], & x>a,\\ 0, & x\le a,\end{cases}$$

где a — параметр сдвига; b — параметр масштаба; c — параметр формы.

Нормально распределенные псевдослучайные числа находились методом обратной функции [38, с. 440, ф-ла (12.10)]. Распределение Вейбулла – Гнеденко моделировалось согласно [39, с. 93].

Вычислительные эксперименты

Ниже приводятся некоторые результаты изучения свойств критериев однородности двух независимых выборок в двух случаях: $F(x)$ и $G(x)$ — функции нормального распределения; $F(x)$ и $G(x)$ — функции распределения Вейбулла – Гнеденко как с одинаковыми, так и с различными значениями параметров.

В первом случае первая выборка бралась из стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1, а вторая — из нормального распределения с математическим ожиданием m_2 и дисперсией σ_2^2, где значения m_2 и σ_2 приведены в табл. 1.

Во втором случае параметр масштаба b функции распределения Вейбулла – Гнеденко во всех выборках принят равным 1. Первая выборка бралась (при всех экспериментах, кроме четырех)

Таблица 1. Проверка равенства математических ожиданий для выборок из нормальных распределений по критерию Крамера – Уэлча

Номер вычислительного эксперимента	Объем выборок $m = n$	Параметры второй выборки		Частота принятия нулевой гипотезы H_0	Вероятность принятия H_0 исходя из распределения Стьюдента	Вероятность принятия H_0 исходя из нормального распределения
		m_2	σ_2			
1	6	0	1	0,969	0,974	0,950
2	7	0	1	0,954	0,971	0,950
3	8	0	1	0,956	0,968	0,950
4	10	0	1	0,958	0,961	0,950
5	6	1	1	0,596	0,691	0,592
6	6	1,5	1	0,366	0,356	0,262
7	8	2	1	0,048	0,032	0,021
8	12	3	1	0	0	0
9	6	0	1,5	0,948	0,974	0,950
10	8	0	2	0,938	0,968	0,950
11	6	0	3	0,930	0,974	0,950
12	10	0	3	0,934	0,961	0,950

при $a = 0$ и $c = 1$, т.е. из экспоненциального распределения с функцией

$$F(x) = \begin{cases} 1 - \exp(-x), x > 0, \\ 0, x \le 0. \end{cases}$$

Вторая выборка бралась из распределений Вейбулла – Гнеденко с параметрами a, $b = 1$, c, приведенными в табл. 2 (там же оговорены исключения).

Выбор распределений для экспериментов определяется желанием как сравнить свойства статистик на выборках из нормального семейства распределений, для которых статистики Стьюдента и Крамера – Уэлча имеют определенные оптимальные свойства, так и рассмотреть класс распределений, существенно отличающихся от нормальных, в частности несимметричностью. Экспоненциальное распределение часто используют при изучении показателей надежности [40], поэтому оно и было включено в эксперименты.

Из семи перечисленных выше критериев однородности критерий Вольфовица (серий), как установлено, имеет малую мощность. Поэтому его исключение из дальнейших рассмотрений не приводит к отрицательным последствиям.

В табл. 3 приведены результаты экспериментов для выборок из нормальных распределений. Табл. 3 соответствует табл. 1 — при совпадающих номерах речь идет об одних и тех же экспериментах. В табл. 4 для облегчения анализа свойств критериев представлены относительные

Таблица 2. Проверка равенства математических ожиданий для выборок из распределений Вейбулла – Гнеденко по критерию Крамера – Уэлча

Номер вычислительного эксперимента	Объем выборок $m = n$	Параметры второй выборки		Частота принятия нулевой гипотезы H_0	Вероятность принятия H_0 исходя из распределения Стьюдента	Вероятность принятия H_0 исходя из нормального распределения
		a	c			
1	6	0	1	0,956	0,974	0,950
2	10	0	1	0,954	0,961	0,950
3	6	0,5	1	0,828	0,912	0,861
4	8	0,5	1	0,772	0,874	0,829
5	10	0,5	1	0,750	0,837	0,800
6	6	1	1	0,750	0,689	0,592
7	8	1	1	0,450	0,558	0,484
8	10	1	1	0,348	0,446	0,313
9	6	0	1,5	0,950	0,971	0,946
10	8	0	1,5	0,950	0,963	0,946
11	10	0	1,5	0,956	0,958	0,942
12	6	0	2	0,940	0,949	0,943
13	8	0	2	0,944	0,954	0,938
14	10	0	2	0,928	0,949	0,935
15	12	0	2	0,944	0,950	0,935
16	8	0	3	0,930	0,961	0,942
17	12	0	3	0,942	0,949	0,935
18	6	0	5	0,904	0,971	0,945
19	8	0	5	0,910	0,963	0,944
20	10	0	5	0,920	0,958	0,943
21	12	0	5	0,940	0,955	0,941
22	8	0	1	0,928	0,968	0,950
23*	6	0	3	0,946	0,974	0,950
24*	10	0	3	0,928	0,961	0,950
25*	6	0,5	3	0,292	0,553	0,447
26*	10	0,5	3	0,094	0,273	0,228
27	6	0,5	3	0,690	0,905	0,850
28	10	0,5	3	0,676	0,826	0,781

* В экспериментах 23 – 26 первая выборка взята из распределения Вейбулла – Гнеденко с параметрами $a = 0$, $b = 1$, $c = 3$.

Таблица 3. Частоты принятия гипотезы однородности для выборок из нормальных распределений

Номер экспери-мента	Объем выборок $m = n$	Параметры второй выборки		Частоты принятия нулевой гипотезы H_0 для критериев					
				1	2	3	5	6	7
		m_2	σ_2	t	L	U	X	S	ω^2
1	6	0	1	0,969	0,970	0,976	0,976	0,982	0,976
2	7	0	1	0,954	0,968	0,986	0,964	1,00	0,956
3	8	0	1	0,956	0,958	0,50	0,954	0,994	0,944
4	10	0	1	0,958	0,960	0,974	0,972	0,998	0,958
5	6	1,0	1	0,596	0,624	0,680	0,698	0,754	0,690
6	6	1,5	1	0,366	0,390	0,464	0,474	0,616	0,496
7	8	2,0	1	0,048	0,054	0,064	0,078	0,480	0,084
8	12	3,0	1	0	0	0	0	0	0
9	6	0	1,5	0,948	0,952	0,974	0,976	0,980	0,972
10	8	0	2,0	0,938	0,940	0,930	0,950	0,998	0,904
11	6	0	3,0	0,930	0,924	0,950	0,956	0,934	0,920
12	10	0	3,0	0,934	0,902	0,930	0,946	0,988	0,846

Таблица 4. Мощность критериев относительно критерия Крамера – Уэлча (для экспериментов № 5 – 12 в табл. 1 и 3)

Номер экспери-мента	Относительная мощность критериев				
	2	3	5	6	7
	L	U	X	S	ω^2
5	0,931	0,792	0,748	0,609	0,767
6	0,727	0,739	0,783	0,000	0,957
7	0,993	0,983	0,968	0,546	0,962
8	1,000	1,000	1,000	1,000	1,000
9	0,923	0,500	0,460	0,385	0,538
10	0,806	1,129	0,806	0,030	1,548
11	1,086	0,714	0,629	0,943	1,140
12	1,485	1,061	0,818	0,182	2,323

мощности критериев по отношению к критерию Крамера – Уэлча (совпадающего с критерием Стьюдента в рассматриваемых экспериментах) — отношения двух случайных величин — оценки мощности рассматриваемого критерия, полученной по 5000 испытаниям, к оценке мощности критерия Крамера – Уэлча.

В табл. 5 приведены результаты экспериментов для выборок из распределений Вейбулла – Гнеденко; табл. 5 соответствует табл. 2 — при совпадающих номерах речь идет об одних и тех же экспериментах.

При анализе табл. 3 и 5 необходимо иметь в виду отличие реальных уровней значимости $\alpha_р$ от номинальных $\alpha_н$ [25]. Особенно это касается критерия Смирнова. Различие между собой реальных уровней значимости у свободных от распределения статистик делает трудным сравнение между собой критериев по мощности — такое сравнение желательно проводить при одном и том же уровне значимости, но это невозможно.

Как и должно быть согласно теории математической статистики [41], для выборок из нормального распределения наиболее мощным оказался критерий Крамера – Уэлча (Стьюдента). Близкими к нему по мощности оказались критерии Лорда и критерий типа омега-квадрат. Критерий Лорда использует размахи, а потому неустойчив к засорениям на «хвостах»; следовательно, возможность его использования при анализе реальных данных в каждом конкретном случае требует специального обоснования. Критерии Вилкоксона и Ван-дер-Вардена также имеют высокую мощность, особенно в экспериментах № 6, 8. Малая мощность критерия Смирнова объясняется, видимо, отличием $\alpha_р$ от $\alpha_н$.

Другая картина наблюдается при изменении дисперсии. Критерии Крамера – Уэлча и Лорда слабо реагируют на нее. Еще меньше реагируют линейные ранговые статистики U и X. Критерий Вилкоксона не может (даже асимптотически) различить нормальные совокупности с одинаковыми математическими ожиданиями, но разными дисперсиями [17]. Обращает на себя внимание высокая мощность критерия омега-квадрат (см. табл. 4, эксперименты № 10 – 12).

Для выборок из распределений Вейбулла – Гнеденко картина несколько иная. Если распределения отличаются только сдвигом (эксперименты № 3 – 8), то наибольшую мощность имеет критерий омега-квадрат, затем идут критерии Вилкоксона и Ван-дер-Вардена, после них — критерии Стьюдента и Лорда, наименьшая мощность у критерия Смирнова. Если же изменяется также и параметр формы (см., например, эксперимент № 21), то наибольшая мощность также у

Таблица 5. Частоты принятия гипотезы однородности для выборок из распределений Вейбулла – Гнеденко

Номер эксперимента	Частоты принятия нулевой гипотезы H_0 для критериев					
	1	2	3	5	6	7
	t	L	U	X	S	ω^2
1	0,956	0,942	0,964	0,968	0,964	0,960
2	0,966	0,940	0,956	0,962	0,998	0,952
3	0,828	0,818	0,840	0,858	0,878	0,840
4	0,772	0,764	0,720	0,746	0,974	0,698
5	0,750	0,724	0,678	0,668	0,962	0,634
6	0,528	0,514	0,534	0,586	0,552	0,482
7	0,450	0,432	0,354	0,392	0,756	0,336
8	0,348	0,344	0,268	0,272	0,662	0,206
9	0,950	0,934	0,958	0,958	0,974	0,958
10	0,950	0,936	0,946	0,954	1,000	0,950
11	0,956	0,934	0,954	0,956	1,000	0,946
12	0,940	0,934	0,962	0,970	0,960	0,952
13	0,944	0,922	0,930	0,972	0,984	0,958
14	0,928	0,900	0,930	0,932	0,990	0,894
15	0,944	0,918	0,938	0,944	0,932	0,898
16	0,930	0,906	0,906	0,924	0,986	0,884
17	0,942	0,908	0,910	0,930	0,816	0,786
18	0,904	0,886	0,934	0,946	0,876	0,866
19	0,910	0,872	0,866	0,896	0,972	0,790
20	0,920	0,872	0,874	0,920	0,968	0,714
21	0,940	0,886	0,862	0,908	0,662	0,606
22	0,928	0,908	0,944	0,952	0,998	0,944
23	0,946	0,948	0,964	0,966	0,978	0,970
24	0,928	0,934	0,948	0,944	0,998	0,936
25	0,292	0,312	0,392	0,408	0,578	0,430
26	0,094	0,100	0,142	0,132	0,654	0,144

критерия омега-квадрат, за ним следует критерий Смирнова (с учетом отличия $\alpha_р$ от $\alpha_н$). Заметно также существенное возрастание мощности с ростом объемов выборок и увеличением различия параметров.

На основе анализа табл. 3 – 5 можно сформулировать (с понятными оговорками) следующие практические рекомендации.

1. Для проверки гипотезы абсолютной однородности (гипотезы совпадения функций распределения двух выборок) целесообразно использовать критерий Лемана – Розенблатта типа омега-квадрат [18] — во всех случаях.

2. Если есть основания предполагать, что распределения отличаются в основном сдвигом, то целесообразно использовать линейные ранговые критерии Вилкоксона и Ван-дер-Вардена. Однако даже в этом случае критерий омега-квадрат может оказаться более мощным.

3. Из рассмотренных критериев для проверки гипотезы однородности в общем случае, кроме критерия ω^2, можно использовать критерий Смирнова — с учетом отличия реального уровня значимости от номинального.

Частота совпадения статистических выводов по разным критериям

По итогам обработки данных с помощью определенного критерия однородности принимают одно из двух решений: «гипотеза однородности отклоняется» или «гипотеза однородности не отклоняется». Решения по разным критериям могут не совпадать. Насколько часты расхождения?

Были изучены доли (в %) расхождений решений по критериям L, U, X, S, ω^2 с решениями по критерию Крамера – Уэлча. Для описания полученных результатов введены «зоны». Пусть t_n — критическое значение для критерия Крамера – Уэлча, соответствующее уровню значимости $\alpha = 0,05$ и объему выборок $m = n$. Используется абсолютное значение статистики критерия Крамера – Уэлча. Введено восемь зон: 1 — $[0; t_n/4)$;

2 — $[t_n/4; t_n/2)$; 3 — $[t_n/2; 3t_n/4)$; 4 — $[3t_n/4; t_n)$; 5 — $[t_n; 5t_n/4)$; 6 — $[5t_n/4; 3t_n/2)$; 7 — $[3t_n/2; 7t_n/4)$; 8 — $[7t_n/4; +\infty)$.

В качестве примера проведенных исследований в табл. 6 представлены данные по вычислительному эксперименту № 19 для выборок из распределений Вейбулла – Гнеденко (см. табл. 2). В строке «T (частота попадания в зону)» приведено асимптотическое распределение статистики Крамера – Уэлча (сгруппированное по зонам). В каждой строке, соответствующей определенному критерию, для каждой зоны указана доля совпадений решений по этому критерию с решением по критерию Крамера – Уэлча.

В качестве следующего примера в табл. 7, построенной аналогично табл. 6, приведена сводка для экспериментов № 1 – 21 с выборками из распределений Вейбулла – Гнеденко. Таблица 8 содержит информацию о расхождениях (в %) решений по критериям L, U, X, S, ω^2 с решениями по критерию Крамера – Уэлча.

Полученные результаты позволяют заключить следующее. Наибольший процент расхождений приходится на зоны 4 (от 10, 4 до 29, 5 % по табл. 7) и 5 (от 6,6 до 77,8 %), что естественно, поскольку при переходе от зоны 4 к зоне 5 и происходит изменение решения по критерию Крамера – Уэлча. Обратим внимание, что расхождения имеются и в зоне 1 — для 6,5 % экспериментов, попавших в эту зону, критерий Лемана – Розенблатта отвергает нулевую гипотезу (т.е. во всех этих случаях гипотеза однородности неверна). Вместе с тем нет ни одного случая, когда бы этот критерий принял гипотезу для экспериментов из зон 7,8. Другими словами, если критерий Крамера – Уэлча отклоняет нулевую гипотезу с $T > 3{,}0$, то критерий ω^2 также отклоняет гипотезу однородности.

Наибольшее расхождение с критерием Крамера – Уэлча наблюдается у критерия Смирнова, в основном за счет принятия гипотезы в случае, когда T-критерий ее отверг. Это во многом объяс-

Таблица 6. Доли совпадений решений по критериям L, U, X, S, ω^2 с решениями по критерию Крамера – Уэлча T (эксперимент № 19)

Критерии	Доля принятия H_0 по T	Зоны							
		1	2	3	4	5	6	7	8
T (частота попадания в зону)	0,910	0,362	0,274	0,194	0,080	0,042	0,022	0,008	0,018
L	0,872	1	1	0,979	0,575	1	1	1	1
U	0,866	0,956	0,978	0,897	0,775	0,562	0,937	1	1
X	0,896	0,978	0,978	0,928	0,900	0,500	0,875	1	1
S	0,972	1	0,985	0,990	1	0,125	0,125	0,750	0,555
ω^2	0,790	0,889	0,927	0,835	0,600	0,875	1	1	1

Таблица 7. Сводка для выборок из распределений Вейбулла – Гнеденко (эксперименты № 1 – 21) — проценты расхождений с решениями по критерию Крамера – Уэлча

Критерии	Зоны							
	1	2	3	4	5	6	7	8
L	0	0	0,9	24,4	6,6	0	0	0
U	1,1	2,2	5,8	22,5	30,7	3,0	0,4	0
X	0,5	1,3	3,5	17,7	34,6	6,2	1,2	0,3
S	3,3	3,4	4,4	10,4	77,8	60,3	46,0	17,0
ω^2	6,5	6,6	11,8	29,5	30,7	2,5	0	0

Таблица 8. Расхождения (в %) решений по критериям L, U, X, S, ω^2 с решениями по критерию Крамера – Уэлча

По критерию Крамера – Уэлча	По другим критериям				
	L	U	X	S	ω^2
Принято 84,6 %, из них отвергнуто	3,2	5,1	3,5	4,5	10,5
Отвергнуто 15,4 %, из них принято	2,7	13,1	15,8	56,5	12,9
Проверено 100 %, из них расхождений	3,1	6,3	5,4	12,4	10,9
По сравнению с критерием Крамера – Уэлча, %	–2,3	–2,3	–0,5	+4,9	–6,9

няется существенным различием $\alpha_р$ и $\alpha_н$ для критерия Смирнова. Почти такое же суммарное число расхождений у критерия Лемана – Розенблатта, но причина иная — у этого критерия мощность выше, чем у критерия Крамера – Уэлча.

Наиболее близок к T-критерию критерий Лорда. Это подтверждается тем, что расхождения имеются лишь в зонах 4 и 5 и незначительное (0,9 %) — в зоне 3.

По числу расхождений критерии Вилкоксона и Ван-дер-Вардена занимают промежуточное положение, они вдвое ближе к статистике Лорда, чем к критериям Смирнова и омега-квадрат. При этом критерий Ван-дер-Вардена ближе к T-критерию, чем критерий Вилкоксона, чего и следовало ожидать, учитывая нацеленность критерия Ван-дер-Вардена на применение к распределениям, близким к нормальным.

При справедливости гипотезы однородности расхождения не превышают 2,2 – 3,2 % и проявляются в зонах 3 – 6. При альтернативе изменения параметра формы расхождения возрастают лишь для критериев Смирнова и ω^2 (до 8,4 – 9,9 %), оставаясь в пределах 3,0 – 4,7 % для остальных критериев, слабо реагирующих на эту альтернативу. При альтернативе сдвига расхождения резко возрастают (до 11,3 % — у критерия Вилкоксона, 10,2 % — у критерия Ван-дер-Вардена, 24,4 % — у критерия Смирнова, 15,8 % — у критерия ω^2), оставаясь малыми (3,7 %) лишь у критерия Лорда.

Можно сделать и ряд других выводов, например, проследить зависимость от объемов выборок и различия параметров. Проведенный нами более детальный анализ подтверждает сформулированные выше практические рекомендации 1 – 3 (завершение раздела «Вычислительные эксперименты»).

Обращает на себя внимание наличие значительного процента расхождений между решениями, принимаемыми по разным критериям. Этот факт необходимо учитывать при обработке конкретных данных в прикладных исследованиях и при разработке нормативно-технической и методической документации, программных продуктов и экспертных систем. В частности, в соответствии с общей теорией устойчивости [42] целесообразно анализировать данные одновременно с помощью нескольких критериев проверки гипотезы однородности двух независимых выборок и затем исходить из выводов, инвариантных относительно выбора критерия.

ЛИТЕРАТУРА

1. **Горский В. Г., Орлов А. И.** Математические методы исследования: итоги и перспективы / Заводская лаборатория. Диагностика материалов. 2002. Т. 68. № 1. С. 108 – 112.
2. **Орлов А. И.** Новая парадигма прикладной статистики / Заводская лаборатория. Диагностика материалов. 2012. Т. 78. № 1. С. 87 – 93.
3. **Орлов А. И.** О новой парадигме математических методов исследования / Научный журнал КубГАУ. 2016. № 122. С. 807 – 832.
4. **Орлов А. И., Луценко Е. В.** Системная нечеткая интервальная математика. — Краснодар: КубГАУ, 2014. — 600 с.
5. **Орлов А. И.** Развитие математических методов исследования (2006 – 2015 гг.) / Заводская лаборатория. Диагностика материалов. 2017. Т. 83. № 1. Ч. 1. С. 78 – 86.
6. **Колмогоров А. Н.** Теория информации и теория алгоритмов. — М.: Наука, 1987. — 304 с.
7. **Григорьев Ю. Д.** Метод Монте-Карло: вопросы точности асимптотических решений и качества генераторов псевдослучайных чисел / Заводская лаборатория. Диагностика материалов. 2016. Т. 82. № 7. С. 72 – 84.
8. **Орлов А. И.** Предельные теоремы и метод Монте-Карло / Заводская лаборатория. Диагностика материалов. 2016. Т. 82. № 7. С. 67 – 72.
9. **Кутузов О. И., Татарникова Т. М.** Из практики применения метода Монте-Карло / Заводская лаборатория. Диагностика материалов. 2017. Т. 83. № 3. С. 65 – 70.
10. **Аронов И. З., Максимова О. В.** Анализ времени достижения консенсуса в работе технических комитетов по стандартизации по результатам статистического моделирования / Заводская лаборатория. Диагностика материалов. 2017. Т. 83. № 3. С. 71 – 77.
11. **Орлов А. И.** Консенсус и истина (комментарий к опубликованной выше статье И. З. Аронова и О. В. Максимовой) / Заводская лаборатория. Диагностика материалов. 2017. Т. 83. № 3. С. 78 – 79.
12. **Орлов А. И.** Значение информационно-коммуникационных технологий для математических методов исследования / Заводская лаборатория. Диагностика материалов. 2017. Т. 83. № 7. С. 5 – 6.
13. **Жуков М. С.** Об алгоритмах расчета медианы Кемени / Заводская лаборатория. Диагностика материалов. 2017. Т. 83. № 7. С. 72 – 78.
14. **Гадолина И. В., Лисаченко Н. Г.** Разработка метода построения доверительных интервалов для процентилей случайной выборки прочности композитов с применением бутстреп-моделирования / Заводская лаборатория. Диагностика материалов. 2017. Т. 83. № 11. С. 73 – 77.
15. **Орлов А. И.** О проверке однородности двух независимых выборок / Заводская лаборатория. Диагностика материалов. 2003. Т. 69. № 1. С. 55 – 60.
16. **Большев Л. Н., Смирнов Н. В.** Таблицы математической статистики. Изд. 3-е. — М.: Наука, 1983. — 416 с.
17. **Орлов А. И.** Какие гипотезы можно проверять с помощью двухвыборочного критерия Вилкоксона? / Заводская лаборатория. Диагностика материалов. 1999. Т. 65. № 1. С. 51 – 55.
18. **Орлов А. И.** Состоятельные критерии проверки абсолютной однородности независимых выборок / Заводская лаборатория. Диагностика материалов. 2012. Т. 78. № 11. С. 66 – 70.
19. **Lehmann E. L.** Consistency and unbiasedness of certain nonparametric tests / Ann. Math. Statist. 1951. Vol. 22. N 2. P. 165 – 179.
20. **Rosenblatt M.** Limit theorems associated with variants of the von Mises statistic / Ann. Math. Statist. 1952. Vol. 23. N 4. P. 617 – 623.
21. **Орлов А. И.** Организационно-экономическое моделирование: учебник. В 3-х ч. Ч. 3. Статистические методы анализа данных. — М.: Изд-во МГТУ им. Н. Э. Баумана, 2012. — 624 с.

22. **Гаек Я., Шидак З.** Теория ранговых критериев. — М.: Наука, 1971. — 374 с.
23. **Холлендер М., Вулф Д.** Непараметрические методы статистики. — М.: Финансы и статистика, 1983. — 520 с.
24. **Парджанадзе А. М., Хмаладзе Э. В.** Об асимптотической теории статистик от последовательных рангов / Теория вероятностей и её применения. 1986. Т. XXXI. Вып. 4. С. 758 – 772.
25. **Орлов А. И.** Реальные и номинальные уровни значимости при проверке статистических гипотез / Научный журнал КубГАУ. 2015. № 114. С. 42 – 54.
26. **Форсайт Дж., Малкольм М., Моулер К.** Машинные методы математических вычислений. — М.: Мир, 1980. — 144 с.
27. **Шеннон Р.** Имитационное моделирование систем: искусство и наука. — М.: Мир, 1978. — 418 с.
28. **Орлов А. И.** Первый Всемирный конгресс Общества математической статистики и теории вероятностей им. Бернулли / Заводская лаборатория. Диагностика материалов. 1987. Т. 53. № 3. С. 90 – 91.
29. **Ермаков С. М.** Метод Монте-Карло и смежные вопросы. — М.: Наука, 1971. — 328 с.
30. **Ермаков С. М., Михайлов Г. А.** Статистическое моделирование. — М.: Наука, 1982. — 296 с.
31. **Журбенко И. Г., Кожевникова И. А., Клиндухова О. В.** Определение критической длины последовательности псевдослучайных чисел / Вероятностно-статистические методы исследования. — М.: МГУ им. М. В. Ломоносова, 1983. С. 18 – 39.
32. **Журбенко И. Г., Кожевникова И. А., Смирнова О. С.** О построении и исследовании псевдослучайных последовательностей различными методами / Заводская лаборатория. Диагностика материалов. 1985. Т. 51. № 5. С. 47 – 51.
33. **Журбенко И. Г.** Анализ стационарных и однородных случайных систем. — М.: МГУ им. М. В. Ломоносова, 1987. — 240 с.
34. **Рыданова Г. В.** Методика изучения временных зависимостей в последовательностях псевдослучайных чисел / Заводская лаборатория. Диагностика материалов. 1986. Т. 52. № 1. С. 56 – 58.
35. **Орлов А. И.** Вероятностно-статистическое моделирование помех, создаваемых электровозами / Научный журнал КубГАУ. 2015. № 106. С. 225 – 238.
36. **Орлов А. И.** Теория люсианов / Научный журнал КубГАУ. 2014. № 101. С. 275 – 304.
37. **Орлов А. И.** О реальных возможностях бутстрепа как статистического метода / Заводская лаборатория. Диагностика материалов. 1987. Т. 53. № 10. С. 82 – 85.
38. **Айвазян С. А., Енюков И. С., Мешалкин Л. Д.** Прикладная статистика. Основы моделирования и первичная обработка данных. — М.: Финансы и статистика, 1983. — 472 с.
39. **Хастингс Н., Пикок Дж.** Справочник по статистическим распределениям. — М.: Статистика, 1980. — 95 с.
40. **Фомин В. Н.** Нормирование показателей надежности. — М.: Изд-во стандартов, 1986. — 140 с.
41. **Кокс Д., Хинкли Д.** Теоретическая статистика. — М.: Мир, 1978. — 560 с.
42. **Орлов А. И.** Устойчивые математические методы и модели / Заводская лаборатория. Диагностика материалов. 2010. Т. 76. № 3. С. 59 – 67.

REFERENCES

1. **Gorskii V. G., Orlov A. I.** Mathematical methods of research: results and prospects / Zavod. Lab. Diagn. Mater. 2002. Vol. 68. N 1. P. 108 – 112 [in Russian].
2. **Orlov A. I.** The new paradigm of Applied Statistics / Zavod. Lab. Diagn. Mater. 2012. Vol. 78. N 1. P. 87 – 93 [in Russian].
3. **Orlov A. I.** On the new paradigm of mathematical research methods / Nauch. Zh. KubGAU. 2016. N 122. P. 807 – 832 [in Russian].
4. **Orlov A. I., Lucenko E. V.** System fuzzy interval mathematics. — Krasnodar: KubGAU, 2014. — 600 p. [in Russian].
5. **Orlov A. I.** Development of the Methods of mathematical Research (2006 – 2015) / Zavod. Lab. Diagn. Mater. 2017. Vol. 83. N 1. Part 1. P. 78 – 86 [in Russian].
6. **Kolmogorov A. N.** Information Theory and Algorithm Theory. — Moscow: Nauka, 1987. — 304 p. [in Russian].
7. **Grigor'ev Yu. D.** Monte Carlo method: accuracy of asymptotic solutions and Quality of Pseudorandom number generators / Zavod. Lab. Diagn. Mater. 2016. Vol. 82. N 7. P. 72 – 84 [in Russian].
8. **Orlov A. I.** Limit theorems and the Monte Carlo method / Zavod. Lab. Diagn. Mater. 2016. Vol. 82. N 7. P. 67 – 72 [in Russian].
9. **Kutuzov O. I., Tatarnikova T. M.** Practical Experience of Using Monte Carlo method / Zavod. Lab. Diagn. Mater. 2017. Vol. 83. N 3. P. 65 – 70 [in Russian].
10. **Aronov I. Z., Maksimova O. V.** Analysis of time to reach consensus on the work of the technical committees of standardization as *a* results of statistical modeling / Zavod. Lab. Diagn. Mater. 2017. Vol. 83. N 3. P. 71 – 77 [in Russian].
11. **Orlov A. I.** Consensus and truth (commens to the article published by I. Z. Aronov and O. V. Maximova) / Zavod. Lab. Diagn. Mater. 2017. Vol. 83. N 3. P. 78 – 79 [in Russian].
12. **Orlov A. I.** Importance of information and communication technologies for mathematical methods of research / Zavod. Lab. Diagn. Mater. 2017. Vol. 83. N 7. P. 5 – 6 [in Russian].
13. **Zhukov M. S.** On the algorithms for Kemeny median calculation / Zavod. Lab. Diagn. Mater. 2017. Vol. 83. N 7. P. 72 – 78 [in Russian].
14. **Gadolina I. V., Lisachenko N. G.** Development of the Method of constructing confidence intervals for percentiles of composites strength random variable using bootstrap simulation / Zavod. Lab. Diagn. Mater. 2017. Vol. 83. N 11. P. 73 – 77 [in Russian].
15. **Orlov A. I.** Test of uniformity of two independent samplings / Zavod. Lab. Diagn. Mater. 2003. Vol. 69. N 1. P. 55 – 60 [in Russian].
16. **Bol'shev L. N., Smirnov N. V.** Tables of mathematical statistics. 3 ed. — Moscow: Nauka, 1983. — 416 p. [in Russian].
17. **Orlov A. I.** Which hypothesis can be tested using the two-sample Wilcoxon Criterion / Zavod. Lab. Diagn. Mater. 1999. Vol. 65. N 1. P. 51 – 55.
18. **Orlov A. I.** Consistent tests of absolute homogeneity of independent sampling / Zavod. Lab. Diagn. Mater. 2012. Vol. 78. N 11. P. 66 – 70.
19. **Lehmann E. L.** Consistency and unbiasedness of certain nonparametric tests / Ann. Math. Statist. 1951. Vol. 22. N 2. P. 165 – 179.
20. **Rosenblatt M.** Limit theorems associated with variants of the von Mises statistic / Ann. Math. Statist. 1952. Vol. 23. N 4. P. 617 – 623.
21. **Orlov A. I.** Organizational and economic modeling: a textbook: in 3 parts. Part 3. Statistical data analysis methods. — Moscow: Izd. MGTU im. N. E. Baumana, 2012. — 624 p. [in Russian].
22. **Hajek Ja., Sidak Zb.** Theory of rank tests. — Prague: Academia. Publishing House of the Czechoslovak Academy of Sciences, 1967. — 376 p.
23. **Hollander M., Wolfe D. A.** Nonparametric statistical methods. — New York – London – Sydney – Toronto: John Wiley and Sons, 1973. — 503 p. [in Russian].
24. **Pardzhanadze A. M., Khmaladze É. V.** On the asymptotic theory of statistics from sequential ranks / Teor. Veroyat. Ee Primen. 1986. Vol. XXXI. N 4. P. 758 – 772 [in Russian].
25. **Orlov A. I.** Real and nominal significance levels in statistical hypothesis testing / Nauch. Zh. KubGAU. 2015. N 114. P. 42 – 54 [in Russian].
26. **Forsythe G. E., Malkolm M. A., Moler C. B.** Computer methods for mathematical computations. — New Jersey: Prentice-Hall, Inc., Englewood Cliffs, 1977. — 192 p.
27. **Shannon R. E.** Systems Simulation: the Art and Science. — New Jersey: Prentice-Hall, Inc., Englewood Cliffs, 1975. — 420 p.

28. **Orlov A. I.** The First World Congress of the Bernoulli Society for Mathematical Statistics and Probability Theory / Zavod. Lab. Diagn. Mater. 1987. Vol. 53. N 3. P. 90 – 91.
29. **Ermakov S. M.** Monte Carlo method and related questions. — Moscow: Nauka, 1971. — 328 p. [in Russian].
30. **Ermakov S. M., Mikhailov G. A.** Statistical modeling. — Moscow: Nauka, 1982. — 296 p. [in Russian].
31. **Zhurbenko I. G., Kozhevnikova I. A., Klindukhova O. V.** Determining the critical length of *a* sequence of pseudo-random numbers / Probabilistic-statistical research methods. — Moscow: MGU im. M. V. Lomonosova, 1983. P. 18 – 39 [in Russian].
32. **Zhurbenko I. G., Kozhevnikova I. A., Smirnova O. S.** On the construction and study of pseudo-random sequences by various methods / Zavod. Lab. Diagn. Mater. 1985. Vol. 51. N 5. P. 47 – 51.
33. **Zhurbenko I. G.** Analysis of stationary and homogeneous random systems. — Moscow: MGU im. M. V. Lomonosova, 1987. — 240 p. [in Russian].
34. **Rydanova G. V.** A technique for studying temporal dependencies in pseudo-random number sequences / Zavod. Lab. Diagn. Mater. 1986. Vol. 52. N 1. P. 56 – 58.
35. **Orlov A. I.** Probabilistic-statistical modeling of the interferences generated by electric locomotives / Nauch. Zh. KubGAU. 2015. N 106. P. 225 – 238.
36. **Orlov A. I.** Theory of lucians / Nauch. Zh. KubGAU. 2014. N 101. P. 275 – 304.
37. **Orlov A. I.** On the real possibilities of bootstrap as *a* statistical method / Zavod. Lab. Diagn. Mater. 1987. Vol. 53. N 10. P. 82 – 85.
38. **Aivazyan S. A., Enyukov I. S., Meshalkin L. D.** Application statistics. Basics of modeling and primary data processing. — Moscow: Finansy i statistika, 1983. — 472 p. [in Russian].
39. **Hastings N. A. J., Peacock J. B.** Statistical distributions. A handbook for students and practitioners. — London: Butterworth and Co (Publishers) Ltd, 1975. — 95 p.
40. **Fomin V. N.** Rationing reliability indicators. — Moscow: Izd. Standartov, 1986. — 140 p. [in Russian].
41. **Cox D. R., Hinkley D. V.** Theoretical statistics. — London: Chapman and Hall, 1974. — 522 p.
42. **Orlov A. I.** Stable mathematical methods and models / Zavod. Lab. Diagn. Mater. 2010. Vol. 76. N 3. P. 59 – 67 [in Russian].