

УДК 519.2:303.732.4

UDC 519.2:303.732.4

**НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ
СОГЛАСИЯ КОЛМОГОРОВА, СМИРНОВА,
ОМЕГА-КВАДРАТ И ОШИБКИ ПРИ ИХ
ПРИМЕНЕНИИ****NONPARAMETRIC GOODNESS-OF-FIT
KOLMOGOROV, SMIRNOV, OMEGA-
SQUARE TESTS AND THE ERRORS IN
THEIR APPLICATION**

Орлов Александр Иванович
д.э.н., д.т.н., к.ф.-м.н., профессор

Orlov Alexander Ivanovich
Dr.Sci.Econ., Dr.Sci.Tech., Cand.Phys-Math.Sci.,
professor
*Bauman Moscow State Technical University,
Moscow, Russia*

*Московский государственный технический
университет им. Н.Э. Баумана, Россия, 105005,
Москва, 2-я Бауманская ул., 5, prof-orlov@mail.ru*

Рассмотрены работы А.Н. Колмогорова и Н.В. Смирнова, посвященные односторонним и двухсторонним критериям согласия и однородности. Показано, что термин «критерий Колмогорова – Смирнова» использовать некорректно. Даны рекомендации по использованию терминов «критерий Колмогорова», «критерий Смирнова», «критерий типа Колмогорова – Смирнова». Обсуждается критерий омега-квадрат (Крамера – Мизеса – Смирнова). Рассмотрены типовые ошибки при применении перечисленных критериев, в частности, для проверки нормальности распределения результатов измерений

In the article we have considered A. N. Kolmogorov and N. V. Smirnov papers dedicated to one-sided and two-sided goodness-of-fit and homogeneity tests. It has been shown that the term "Kolmogorov - Smirnov test" used incorrectly. We have also given the recommendations on use of the terms "Kolmogorov test", "Smirnov test", "test of Kolmogorov-Smirnov type" and discussed omega-square test (Cramer-von Mises–Smirnov test). Typical errors in the application of these criteria have been considered, in particular to test for normality of the distribution of measurement results

Ключевые слова: МАТЕМАТИЧЕСКАЯ СТАТИСТИКА, КРИТЕРИИ СОГЛАСИЯ, КРИТЕРИИ КОЛМОГОРОВА, КРИТЕРИИ СМИРНОВА, КРИТЕРИИ ОМЕГА-КВАДРАТ, ПРЕДЕЛЬНЫЕ РАСПРЕДЕЛЕНИЯ, ОШИБКИ ПРИ АНАЛИЗЕ ДАННЫХ, ИСТОРИЯ СТАТИСТИКИ

Keywords: MATHEMATICAL STATISTICS, GOODNESS-OF-FIT TEST, KOLMOGOROV TEST, SMIRNOV TESTS, OMEGA-SQUARE TEST, LIMIT DISTRIBUTIONS, ERRORS IN DATA ANALYSIS, HISTORY OF STATISTICS

1. Введение

Непараметрические критерии согласия Колмогорова, Смирнова, омега-квадрат, разработанные в 1930-е годы, являются основой непараметрической математической статистики. Они продолжают быть образцом при разработке новых методов и широко применяются на практике. Используются термины «статистики типа Колмогорова – Смирнова», «статистики типа омега-квадрат». Встречается в литературе и в описании программных продуктов некорректный термин «статистика Колмогорова – Смирнова». А также и другие некорректности – например, пишут о применении критерия Колмогорова для проверки согласия с семейством нормальных распределений. В последнем из упомянутых случаев речь отнюдь не только о терминологии – часты грубые ошибки

при применении непараметрических критериев математической статистики, приводящие к неправильным управленческим решениям.

В настоящей статье рассмотрена система непараметрических критериев согласия Колмогорова, Смирнова, омега-квадрат и им аналогичных, а также проанализированы типовые ошибки при применении этих критериев.

Введем эмпирическую функцию распределения - основное понятие, используемое в статье. Рассмотрим конечную последовательность независимых одинаково распределенных числовых случайных величин, т.е. выборку $\xi_1, \xi_2, \dots, \xi_n$, где n – объем выборки, а $F(x) = P(\xi_i < x)$ - функция распределения элементов выборки. Эмпирической функцией распределения называется функция $F_n(x)$, для каждого действительного числа x равная доле наблюдений, меньших x , т.е.

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \chi(x - \xi_i),$$

где $\chi(u) = 1$ при $u > 0$ и $\chi(u) = 0$ при $u \leq 0$. Эмпирическая функция распределения $F_n(x)$ является состоятельной оценкой теоретической функции распределения $F(x)$, т.е. функции распределения элементов выборки. Справедлива теорема В.И. Гливенко: для любой непрерывной функции распределения $F(x)$ при $n \rightarrow \infty$ эмпирическая функция распределения $F_n(x)$ равномерно сходится к ней:

$$\sup_x |F_n(x) - F(x)| \rightarrow 0$$

(сходимость по вероятности).

2. О критериях Колмогорова и Смирнова

В литературе, особенно переводной, иногда используют термин «критерий Колмогорова – Смирнова» по отношению к процедурам проверки непараметрических статистических гипотез. Однако анализ публикаций академика А.Н. Колмогорова (1903 – 1987) и член-

корреспондента АН СССР Н.В. Смирнова (1900 – 1966), проведенный ниже, свидетельствует о том, что такого критерия **не существует**. У А.Н. Колмогорова и Н.В. Смирнова нет совместных работ, они никогда не изучали одновременно один и тот же статистический критерий. Право на существование имеет лишь термин «критерий типа Колмогорова – Смирнова».

Научное наследие А.Н. Колмогорова и Н.В. Смирнова отнюдь не потеряло актуальности. Например, в статье «Метод медиан в теории ошибок» 1931 г. [1] А.Н. Колмогоров показывает преимущество выборочной медианы результатов измерений перед выборочным средним арифметическим при оценивании истинного значения контролируемого параметра, если погрешности измерений имеют те или иные распределения. В работе «К вопросу о пригодности найденных статистическим путем формул прогноза» 1933 г. [2] он обсуждает методы прогнозирования на основе отбора информативных признаков в регрессии, указывает на типичные ошибки при интерпретации расчетов. Эти ошибки делаются и поныне. В работе Н.В. Смирнова [3] 1947 г. предложен и изучен непараметрический критерий «типа Колмогорова – Смирнова» для проверки симметрии функции распределения случайной величины относительно 0, который в настоящее время целесообразно включать в пакеты программ по прикладной статистике (как и критерий типа омега-квадрат для проверки той же гипотезы, предложенный и изученный в работе [4], через 32 года адаптированный для практического использования в статье [5]).

Рассмотрим работы А.Н. Колмогорова и Н.В. Смирнова, связанные с «критериями типа Колмогорова – Смирнова». (Все упомянутые в статье работы этих авторов перепечатаны в сборниках трудов [6, 7]).

В 1933 г. была опубликована на итальянском языке знаменитая статья А.Н. Колмогорова «Об эмпирическом определении закона распределения» [8], в которой была доказана следующая теорема:

«**Теорема 1.** Вероятность $\Phi_n(\lambda)$ неравенства

$$D = D_n = \sup_x |F_n(x) - F(x)| < \lambda / \sqrt{n} \quad (1)$$

стремится при $n \rightarrow \infty$ равномерно по λ к

$$\Phi(\lambda) = \sum_{-\infty}^{\infty} (-1)^k \exp(-2k^2 \lambda^2) \quad (2)$$

для любой непрерывной функции распределения $F(x)$ ».

Здесь $F_n(x)$ – эмпирическая функция распределения, построенная по выборке объема n из непрерывной функции распределения $F(x)$, т.е. $nF_n(x)$ – число элементов выборки, не превосходящих x .

Итак, в работе [8] найдено, выражаясь современным языком, предельное распределение статистики критерия Колмогорова для проверки согласия эмпирического распределения с заданным непрерывным теоретическим распределением.

В 1940 г. опубликована статья А.Н. Колмогорова «Об одном новом подтверждении законов Менделя» [9]. Он проанализировал статистические данные Н.И. Ермолаевой, считавшей, что эти данные опровергают законы Менделя. С помощью критерия Колмогорова в статье [9] было показано, что эти данные, наоборот, полностью подтверждают законы Менделя. Подробный анализ ситуации дан в работе [10].

Таким образом, к рассматриваемой тематике относятся две работы А.Н. Колмогорова: в первой из них он ввел свой знаменитый критерий для проверки согласия эмпирического распределения с заданным теоретическим и нашел предельное распределение статистики критерия, а во второй – применил его для решения важной научной задачи в области биологии. Больше к этой тематике он не возвращался (см. итоговый сборник статей [6]).

В те же 1930-е годы Н.В. Смирнов изучал различные непараметрические статистики. В статье 1939 г. «Об уклонениях эмпирической кривой распределения» [11] он рассматривал число пересечений эмпирической функции распределения $F_n(x)$ с кривой $F(x) + \lambda/\sqrt{n}$ и число выходов $F_n(x)$ за пределы полосы $F(x) \pm \lambda/\sqrt{n}$ при произвольном фиксированном $\lambda > 0$.

Из предельной теоремы для числа пересечений эмпирической функции распределения с указанной выше кривой было получено утверждение об асимптотическом поведении статистики (в настоящее время ее называют «односторонней одновыборочной статистикой Смирнова»)

$$\tau_n = \sup_x (F_n(x) - F(x)). \quad (3)$$

А именно, в статье [11] сказано: «На основании теоремы II мы можем утверждать для $n \rightarrow \infty$:

$$P\left\{\tau_n \leq \frac{\lambda}{\sqrt{n}}\right\} \rightarrow 1 - \exp(-2\lambda^2), \quad \lambda \geq 0. \quad (4)$$

Аналогично частным случаем теоремы о предельном поведении числа выходов эмпирической функции распределения за пределы полосы является, согласно статье [11], указанная выше теорема 1 А.Н. Колмогорова [8]. Подчеркнем, что методы доказательства в работах А.Н. Колмогорова [8] и Н.В. Смирнова [11] принципиально различны.

В опубликованной в том же 1939 г. статье «Оценка расхождения между эмпирическими кривыми распределения в двух независимых выборках» [12] Н.В. Смирнов изучает статистику (в настоящее время ее называют «двухсторонней двухвыборочной статистикой Смирнова»)

$$D(n(1), n(2)) = \sup_{-\infty < x < \infty} |F_{n(1)}(x) - F_{n(2)}(x)|, \quad (5)$$

где $F_{n(1)}(x)$ – эмпирическая функция распределения, построенная по первой выборке объема $n(1)$ из непрерывной функции распределения $F_1(x)$, а

$F_{n(2)}(x)$ – эмпирическая функция распределения, построенная по второй выборке объема $n(2)$ из непрерывной функции распределения $F_2(x)$, независимой от первой выборки. Статистика $D(n(1), n(2))$ используется для проверки гипотезы однородности двух независимых выборок, т.е. гипотезы

$$H_0 : F_1(x) \equiv F_2(x). \quad (6)$$

В случае справедливости этой гипотезы в работе [12] доказана следующая теорема:

«**Теорема 1.** Если объемы выборок $n(1)$ и $n(2)$ неограниченно возрастают так, что отношение $n(1)/n(2) = \tau$ остается постоянным, то

$$P_{n(1),n(2)}(\lambda) = P\left(D(n(1), n(2)) \leq \frac{\lambda}{\sqrt{n}}\right) \rightarrow \Phi(\lambda), \quad (7)$$

где

$$\lambda > 0, n = \frac{n(1)n(2)}{n(1) + n(2)} = \frac{n(2)}{1 + \tau}, \quad (8)$$

а $\Phi(\lambda)$ - функция Колмогорова», заданная формулой (2).

Как и в работе [11], Н.В. Смирнов выводит эту теорему из более общего утверждения, касающегося числа точек пересечения кривых $F_{n(1)}(x)$ и $F_{n(2)}(x) \pm \lambda/\sqrt{n}$.

В статье [12] были приведены также таблицы функции Колмогорова $\Phi(\lambda)$. Наиболее подробные таблицы $\Phi(\lambda)$, рассчитанные в 1965 г. под руководством Н.В. Смирнова, опубликованы в работе [7, с.267-277].

Следует отметить, что условие $n(1) \rightarrow \infty, n(2) \rightarrow \infty, n(1)/n(2) = \tau$ в дальнейшем было заменено на более слабое: $\min(n(1), n(2)) \rightarrow \infty$ (см., например, монографию [13]). Математический аппарат, с помощью которого можно обосновать указанную замену, рассмотрен в разделе 2.6 монографии [14] (о двухсторонней двухвыборочной статистике Смирнова см. с.241).

В той же статье [12] Н.В. Смирнова найдено асимптотическое распределение односторонней статистики в случае двух выборок (в настоящее время ее называют «односторонней двухвыборочной статистикой Смирнова»)

$$D^+(n(1), n(2)) = \sup_{-\infty < x < \infty} (F_{n(1)}(x) - F_{n(2)}(x)).$$

Позже [15] Н.В. Смирновым найдено точное распределение одновыборочной односторонней статистики τ_n [см. формулу (3)], из которого им выведено асимптотическое распределение τ_n . Это – другой подход к получению соотношения (4), чем развитый в статье [11]. В качестве следствий Н.В. Смирнов получает теорему Гливленко и теоремы типа закона повторного логарифма для статистики Колмогорова, а также теорему 1 Колмогорова, приведенную выше.

Рассмотренные работы А.Н. Колмогорова и Н.В. Смирнова положили начало большому научному направлению в рамках непараметрической математической статистики. Если А.Н. Колмогоров, опубликовав две статьи, переключился на другие задачи, то Н.В. Смирнов продолжал развивать и применять свой виртуозный математический аппарат. В частности, им предложены [3] критерии для проверки симметрии распределения относительно 0 в стиле статистик D и τ_n . Математический аппарат, развитый Н.В. Смирновым для анализа распределений приведенных выше статистик, оказался полезным и для изучения непараметрических оценок плотности [16], так что Н.В. Смирнова следует считать основоположником и этого направления математической статистики.

Появились варианты статистики Смирнова τ_n , в которых супремум брался не по всей прямой, а по некоторому отрезку, или же разность делилась на $F(x)$, т.е. бралась в относительном виде. Эти статистики изучали Г.М. Мания, А. Реньи, В.С. Королюк, Чжан Ли-цян, Горо Ишии и

др. Позднее в работе [17] Н.В. Смирнов описал точные распределения, а также асимптотику вероятностей больших значений для всех этих статистик.

Основной книгой по прикладной математической статистике на русском языке являются «Таблицы математической статистики» Л.Н. Большева и Н.В. Смирнова [18]. Кроме непосредственно таблиц, в этой книге имеется большой справочный раздел по основным статистическим методам. В главе «Таблицы непараметрической статистики» дана сводка по состоянию на 1965 г. – год первого издания книги. В отличие от других справочных изданий, данная книга практически не содержит ошибок.

Исследования в рассматриваемой области непараметрической статистики продолжаются (см., например, [13],[18-21]). Критерии D , τ_n , $D(n(1), n(2))$, $D^+(n(1), n(2))$ рассматриваются в учебных курсах, включаются в пакеты статистических программ. Поэтому нельзя не обращать внимания на их наименования.

Проведенный выше анализ публикаций А.Н. Колмогорова и Н.В. Смирнова позволяет сделать следующие заключения.

1. Двухвыборочные критерии, основанные на статистиках $D(n(1), n(2))$ и $D^+(n(1), n(2))$, должны носить имя Н.В. Смирнова.

2. Одновыборочные односторонние критерии, т.е. критерий на основе статистики τ_n и аналогичный критерий, в котором вместо супремума берется инфимум, - это критерии Смирнова.

3. Одновыборочный двухсторонний критерий D – это критерий Колмогорова.

Заключения 1 и 2 очевидны – А.Н. Колмогоров не занимался этими критериями, они введены и изучены Н.В. Смирновым. Заключение 3 также вряд ли вызовет возражения: хотя Н.В. Смирнов передоказывал теорему А.Н. Колмогорова, изучал свойства критерия D , составлял таблицы предельного распределения, А.Н. Колмогоров раньше получил все

результаты, необходимые для практического использования критерия D , в том числе в работе [8] приведена первая таблица функции $\Phi(\lambda)$.

Разумеется, все рассматриваемые в настоящем разделе критерии – это критерии «типа Колмогорова-Смирнова». Что же это за «тип»? Так обычно называют критерии, построенные на основе эмпирических функций распределения (для одной или нескольких выборок), в которых используются операция взятия супремума. Примерами являются статистики, предложенные и изученные Г.М. Манья, А. Реньи, В.С. Королук, Чжан Ли-цян, Горо Ишии и др., упомянутые выше в связи с работой [17]. Один из вариантов строгой формулировки понятия «критерий типа Колмогорова – Смирнова» дан в работе [22].

Почему же в англоязычной литературе (см., например, [23]), а под влиянием ее недостаточно квалифицированных переводов и на русском языке двухвыборочный критерий проверки однородности [т.е. гипотезы (6)], основанный на статистике $D(n(1), n(2))$, называют «критерием Колмогорова – Смирнова»? В работе [23, с.239-240] сказано: «Смирнов получил предельное распределение статистики Колмогорова Смирнова¹, основываясь на работе Колмогорова о предельном распределении одновыборочной статистики D ». Ясно, что авторы цитируемой работы не читали соответствующей работы Н.В. Смирнова [12], иначе они бы знали, что метод доказательства в работе [12] не имеет ничего общего с методом доказательства в статье А.Н. Колмогорова [8]. Нет ничего удивительного в том, что не читали, поскольку работы как [8], так и [12] были малодоступны для англоязычных статистиков (статья А.Н. Колмогорова [8] была напечатана на итальянском языке, а статья Н.В. Смирнова [12] – на русском).

Теперь, когда все эти работы стали широко доступны российским специалистам (см. [6] и [7]), пора отбросить старые недоразумения и

¹ Речь идет о статистике $D(n(1), n(2))$.

воздать должное великим статистикам А.Н. Колмогорову и Н.В. Смирнову.

Подчеркнем, что применение термина «критерий Колмогорова – Смирнова» к двухвыборочному критерию проверки однородности [т.е. гипотезы (6)], основанный на статистике $D(n(1), n(2))$, свидетельствует о том, что автор соответствующего текста не знаком с работами А.Н. Колмогорова и Н.В. Смирнова, пишет о том, чего не знает, т.е. является невеждой и вводит читателя в заблуждение.

В статье [24] показано, что термин «критерий Колмогорова – Смирнова» применительно к критериям на основе статистик D , τ_n , $D(n(1), n(2))$, $D^+(n(1), n(2))$ использовать некорректно, и даны рекомендации по адекватному употреблению терминов «критерий Колмогорова», «критерий Смирнова», «критерий типа Колмогорова – Смирнова».

3. Проверка согласия эмпирического распределения с заданным теоретическим

Критерий Колмогорова предназначен для проверки согласия эмпирического распределения с заданным непрерывным теоретическим распределением, т.е. для проверки гипотезы

$$H_0 : F(x) = F_0(x) \quad (9)$$

при всех возможных значениях аргумента x , где $F(x)$ – функция распределения элементов выборки, $F_0(x)$ – заданная (т.е. полностью известная) функция теоретического распределения. Нулевую гипотезу (9) можно записать в виде

$$\sup_x |F(x) - F_0(x)| = 0. \quad (10)$$

Альтернативная гипотеза состоит в отрицании нулевой гипотезы, т.е.

$$\sup_x |F(x) - F_0(x)| > 0. \quad (11)$$

Для проверки гипотезы (9) – (10) выбирают уровень значимости α , рассчитывают значение статистики

$$\sqrt{n}D_n = \sqrt{n} \sup_x |F_n(x) - F_0(x)|,$$

при

$$\sqrt{n}D_n \leq C(\alpha)$$

закljučают, что нулевая гипотеза (9) – (10) не отклоняется (как обычно несколько неточно говорят, принимается), а при

$$\sqrt{n}D_n > C(\alpha)$$

отклоняется нулевая гипотеза (9) – (10) в пользу альтернативной гипотезы (11) (проще говоря, принимается альтернативная гипотеза). Граничное значение $C(\alpha)$ - квантиль функции распределения $\Phi(\lambda)$,

$$\Phi(C(\alpha)) = 1 - \alpha. \quad (12)$$

Граничные значения $C(\alpha)$ для наиболее распространенных уровней значимости α приведены в табл.1.

Для проверки нулевой гипотезы (9) используют также критерий омега-квадрат. В работах Г. Крамера (1928) [25], Р. фон Мизеса (1931) [26] и Н.В. Смирнова (1936, 1937) [27, 28] (см. также [15]) была предложена статистика

$$\omega_n^2 = n \int_{-\infty}^{\infty} (F_n(x) - F_0(x))^2 dF_0(x), \quad (13)$$

которая теперь называется статистикой Крамера – Мизеса – Смирнова. Отметим роль В.И. Гливленко, который предложил заменить интегрирование по dx , как было в первоначальном варианте, на интегрирование по $dF_0(x)$. Такая модификация привела к тому, что при справедливости нулевой гипотезы (9) распределение статистики ω_n^2 не зависит от вида непрерывной функции $F_0(x)$.

В работах [27, 28] Н.В. Смирнов показал, что функция распределения $S_n(x)$ статистики ω_n^2 сходится при $n \rightarrow \infty$ к некоторой предельной функции распределения $S_0(x)$, а именно,

$$S_0(x) = 1 - \frac{2}{\pi} \sum_{k=0}^{\infty} (-1)^k \int_{(2k-1)\pi}^{2k\pi} \frac{e^{-\frac{\lambda^2 x}{2}}}{\sqrt{-\lambda \sin \lambda}} d\lambda. \quad (14)$$

Отметим, что в работах [27, 28] множитель $(-1)^k$ был пропущен. Этот недостаток был исправлен нами в [29] (совместно с С.А. Пироговым).

В [13, разд.2.3] дан обзор результатов, касающихся следующего шага исследований – оценки скорости сходимости $S_n(x)$ к $S_0(x)$.

Нулевую гипотезу (9) можно записать в виде

$$\int_{-\infty}^{\infty} (F(x) - F_0(x))^2 dF_0(x) = 0. \quad (15)$$

Альтернативная гипотеза состоит в отрицании нулевой гипотезы, т.е.

$$\int_{-\infty}^{\infty} (F(x) - F_0(x))^2 dF_0(x) > 0. \quad (16)$$

Для проверки гипотезы (9) – (15) выбирают уровень значимости α , рассчитывают значение статистики ω_n^2 , при

$$\omega_n^2 \leq B(\alpha)$$

закljučают, что нулевая гипотеза (9) – (15) не отклоняется, а при

$$\omega_n^2 > B(\alpha)$$

принимается альтернативная гипотеза (16). Граничное значение $B(\alpha)$ – квантиль функции распределения $S_0(x)$:

$$S_0(B(\alpha)) = 1 - \alpha.$$

Граничные значения $B(\alpha)$ для наиболее распространенных уровней значимости α приведены в табл.2.

По образцу статистики (13) разработаны многочисленные статистики интегрального типа [30, 31], предназначенные для проверки различных гипотез. Часто их можно назвать «статистики типа омега-квадрат».

Например, в [4] сконструирована статистика типа омега-квадрат для проверки симметрии функции распределения относительно 0.

Выше описаны правила проверки статистических гипотез с использованием предельных распределений соответствующих статистик. Для малых объемов выборок рассчитаны таблицы, позволяющие принимать обоснованные решения [18, 23].

4. О критериях согласия с параметрическим семейством

Непараметрические критерии согласия Колмогорова, Смирнова, омега-квадрат широко используются на практике, включены в программные продукты и в учебные материалы. Однако с ними связаны и широко распространенные ошибки в применении статистических методов.

Дело в том, что перечисленные критерии были разработаны для проверки согласия с *полностью известным* теоретическим распределением. Расчетные формулы, таблицы распределений и критических значений широко распространены, поэтому не будем их приводить. В частности, они имеются в наилучшем на наш взгляд сборнике статистических таблиц Л.Н. Большева и Н.В.Смирнова [18].

Основная идея критериев Колмогорова, омега-квадрат и аналогичных им [18] состоит в измерении расстояния между функцией эмпирического распределения и функцией теоретического распределения. Различаются эти критерии видом расстояний в пространстве функций распределения.

Часто возникает задача проверки гипотезы согласия эмпирического распределения с параметрическим семейством, например, с семейством нормальных, Вейбулла-Гнеденко или гамма-распределений. Профанам представляется естественным оценить параметры распределения по выборке, а затем формально воспользоваться критериями согласия Колмогорова, Смирнова или омега-квадрат. При этом вместо

фиксированной теоретической функции распределения подставляют функцию из параметрического семейства, в которой параметры заменены на их выборочные оценки.

В отличие от классических критериев, при этом измеряются расстояния от эмпирической функции распределения до многообразий (в пространстве функций распределения), соответствующих параметрическим семействам. Развита математическая техника проектирования в функциональных пространствах [32], которая позволяет строить методы проверки и иных гипотез, в частности, гипотезы симметрии функции распределения относительно 0 [4].

Такие критерии согласия уже не являются классическими. В частности, их нельзя считать непараметрическими. Их распределения (как предельные, так и при конечных объемах выборок) *существенно отличаются* от распределений классических критериев согласия Колмогорова, Смирнова или омега-квадрат. Такие критерии в отличие от классических обычно называют «критериями согласия с параметрическим семейством типа Колмогорова-Смирнова и типа омега-квадрат». (Как показано выше (и в [24]) на основе анализа исходных публикаций, термин «критерий Колмогорова-Смирнова» употреблять некорректно.)

Широко распространенные ошибки состоят в том, что для критериев согласия с параметрическими семействами используют критические значения классических критериев. При этом, например, гипотеза нормальности принимается гораздо чаще, чем следует. Поскольку в действующей нормативно-технической документации дальнейшие этапы анализа данных часто зависят от того, принимается нормальность или нет, то ошибки при такой проверке могут иметь далеко идущие последствия. Приведем пример из нашего опыта: при анализе характеристик эластомерных материалов при ошибочном подходе из 30 выборок

нормальность была отвергнута лишь для 2, а при правильном - для 26, т.е. в подавляющем большинстве случаев [33].

Указанные ошибки встречаются в массе публикаций (хотя специалистам суть дела хорошо известна уже около 60 лет, со статьи Каца, Кифера и Вольфовица [34] 1955 г.). Наиболее известным примером является полностью ошибочный ГОСТ 11.006-74 (СТ СЭВ 1190-78) «Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим» (отменен в 1987 г. после яростного сопротивления разработчиков этого шедевра А.М. Бендерского и А.А. Богатырева; о большом числе ошибочных стандартов по статистическим методам рассказывается в [35, 36]).

В консультации [19] собраны основные факты о критериях согласия с параметрическими семействами типа Колмогорова-Смирнова и типа омега-квадрат и необходимые краткие таблицы. Современное положение дел в этой области отражено в дальнейших публикациях в журнале «Заводская лаборатория» [20, 21, 37].

Наибольшее продвижение в изучении критериев типа Колмогорова-Смирнова достигнуто Ю.Н.Тюриным и его ученицами [21, 38-40]. В частности, найдены удобные формулы для асимптотического достигаемого уровня значимости.

Исчерпывающие результаты по расчету предельных распределений критериев типа омега-квадрат приведены в монографии Г.В. Мартынова [41]. Скорость сходимости распределений критериев при конечных объемах выборки к предельным распределениям [18] изучалась теоретически в [29] и численно в [37]. Оказалось, в частности, что предельным распределением можно пользоваться при объеме выборки не менее 4, если уровень значимости не меньше 0,01.

В [20] описан оригинальный подход к рассматриваемой задаче, принадлежащий Дж. Дурбину [42]. Оказывается, если в теоретическую

функцию распределения подставлять оценку максимального правдоподобия (или ей эквивалентную, например, одношаговую оценку [43]), построенную не по всей выборке, а по ее половине, то предельные распределения статистик типа Колмогорова-Смирнова и типа омега-квадрат будут совпадать с распределениями соответствующих классических статистик!

Для проверки согласия с параметрическим семейством используют также критерий хи-квадрат, при применении которого распространены свои типичные ошибки, разобранные М. Мирвалиевым и М.С. Никулиным [44] и В.Г. Воиновым [45]. Не потерял своего значения для проверки согласия и метод моментов [46].

Популярны и другие ошибки при применении рассматриваемых критериев. Некоторые пытаются их использовать для сгруппированных данных, что приводит к излишне частому принятию гипотез (см. монографию Р. Руниона [47]). Другие вместо эмпирической функции распределения рассматривают иные оценки теоретической функции распределения. Например, при использовании вероятностной бумаги удобно ординату точки, соответствующей i -ой порядковой статистике, установить равной $(i - 0.5) / n$, а не i / n , как в классической эмпирической функции распределения. Возникает искушение построенную так оценку использовать в критериях согласия вместо эмпирической функции распределения. Увы, распределение изменится (впрочем, в данном случае при росте объема выборки различие будет исчезать). Ряд иных ошибок рассмотрен в [18].

В целом, по нашему мнению, рассматриваемая тематика принадлежит к тупиковой линии развития. Из общих соображений очевидно, что распределения реальных данных никогда не входят в какое-либо параметрическое семейство. Другими словами, эти семейства - лишь возможные приближения, заведомо неточные. Анализ конкретных данных

приводит к аналогичному заключению. Так, в статье [48] собраны факты, показывающие, что распределения ошибок измерений почти всегда отличаются от нормальных.

Следовательно, от методов параметрической статистики целесообразно переходить к непараметрическим и робастным методам. Однако это не значит, что, применяя параметрические методы, в частности, проверяя согласие с параметрическим семейством распределений, можно делать математические ошибки типа описанных выше. Приходится с сожалением констатировать, что распространенные программные продукты по статистическим методам анализа данных, как правило, отстали от современного уровня развития науки лет на 30 и зачастую содержат ошибки, в том числе описанные выше [49].

5. О проверке нормальности с помощью критериев типа Колмогорова — Смирнова и типа омега-квадрат

Непараметрическая статистика включает в себя не только проверку гипотез с помощью рассмотренных выше критериев. В качестве примера непараметрического оценивания рассмотрим оценивание функции распределения. По теореме Гливленко эмпирическая функция распределения $F_n(x)$ является состоятельной оценкой функции распределения $F(x)$. Если $F(x)$ – непрерывная функция, то на основе теоремы Колмогорова доверительные границы для функции распределения $F(x)$ задают в виде

$$(F(x))_{\text{нижн}} = \max\left\{0, F_n(x) - \frac{k(\gamma, n)}{\sqrt{n}}\right\}, \quad (F(x))_{\text{верх}} = \min\left\{1, F_n(x) + \frac{k(\gamma, n)}{\sqrt{n}}\right\},$$

где $k(\gamma, n)$ – квантиль порядка γ распределения статистики Колмогорова при объеме выборки n (напомним, что распределение этой статистики не зависит от $F(x)$).

Как известно [50], правила определения оценок и доверительных границ в параметрическом случае строятся на основе параметрического семейства распределений $F(x; \theta)$. При обработке реальных данных возникает вопрос – соответствуют ли эти данные принятой вероятностной модели? То есть статистической гипотезе: результаты наблюдений имеют функцию распределения из семейства $\{F(x; \theta), \theta \in \Theta\}$ при некотором $\theta = \theta_0$. Такие гипотезы называют гипотезами согласия, а критерии их проверки – критериями согласия.

Если истинное значение параметра $\theta = \theta_0$ известно, функция распределения $F(x; \theta_0)$ непрерывна, то для проверки гипотезы согласия часто применяют критерий Колмогорова, основанный на статистике

$$\sqrt{n}D_n = \sqrt{n} \sup_x |F_n(x) - F(x; \theta_0)|$$

где $F_n(x)$ – эмпирическая функция распределения, $F(x; \theta_0) = F_0(x)$ – полностью известная теоретическая функция распределения.

Если истинное значение параметра θ_0 неизвестно, например, при проверке гипотезы о нормальности распределения результатов наблюдения (т.е. при проверке принадлежности этого распределения к семейству нормальных распределений), то иногда используют статистику типа Колмогорова-Смирнова (или, как выражаются короче, типа Колмогорова)

$$\sqrt{n}D_n(\theta^*) = \sqrt{n} \sup_x |F_n(x) - F(x; \theta^*)|.$$

Она отличается от статистики Колмогорова $\sqrt{n}D_n$ тем, что вместо истинного значения параметра θ_0 подставлена его оценка θ^* .

Распределение статистики $\sqrt{n}D_n(\theta^*)$ сильно отличается от распределения статистики D_n . В качестве примера рассмотрим проверку нормальности, когда $\theta = (m, \sigma^2)$, а $\theta^* = (\bar{x}, s^2)$ (т.е. по выборке оцениваются как математическое ожидание, так и дисперсия нормального распределения). Для этого случая квантили распределений статистик $\sqrt{n}D_n$

и $\sqrt{n} D_n(\theta^*)$ приведены в табл.1 (см., например, [19]). Таким образом, квантили отличаются примерно в 1,5 раза. Для приведенных в нижней строке табл.1 уровней значимости соответствующие квантили являются критическими значениями при проверке нулевых гипотез, описанных выше.

Таблица 1. Квантили статистик Колмогорова $\sqrt{n} D_n$ и типа Колмогорова – Смирнова $\sqrt{n} D_n(\theta^*)$ при проверке нормальности

Порядок p квантиля	0,85	0,90	0,95	0,975	0,99
Квантили порядка p для $\sqrt{n} D_n$	1,138	1,224	1,358	1,480	1,626
Квантили порядка p для $\sqrt{n} D_n(\theta^*)$	0,775	0,819	0,895	0,955	1,035
Уровень значимости α	0,15	0,10	0,05	0,025	0,01

В ряде литературных источников и программных продуктов при описании и использовании критериев согласия допускается грубая ошибка: при проверке нормальности вместо критических значений статистики $\sqrt{n} D_n(\theta^*)$ применяют критические значения статистики $\sqrt{n} D_n$. В результате гипотеза нормальности принимается гораздо чаще, чем следует. Эта ошибка подробно рассмотрена в статье [19].

Аналогична ситуация с неправомерным использованием статистики омега-квадрат. Если истинное значение параметра $\theta = \theta_0$ известно, функция распределения $F(x; \theta_0)$ непрерывна, то для проверки гипотезы согласия применяют критерий омега-квадрат (Крамера – Мизеса – Смирнова), основанный на статистике

$$\omega_n^2(\theta_0) = n \int_{-\infty}^{\infty} (F_n(x) - F(x; \theta_0))^2 dF(x; \theta_0),$$

где $F_n(x)$ – эмпирическая функция распределения, $F(x; \theta_0) = F_0(x)$ – полностью известная статистику теоретическая функция распределения.

Если истинное значение параметра θ_0 неизвестно, например, при проверке гипотезы о нормальности распределения результатов наблюдения (т.е. при проверке принадлежности этого распределения к семейству нормальных распределений), то иногда используют статистику типа омега-квадрат

$$\omega_n^2(\theta^*) = n \int_{-\infty}^{\infty} (F_n(x) - F(x; \theta^*))^2 dF(x; \theta^*).$$

Она отличается от статистики омега-квадрат (Крамера – Мизеса – Смирнова) $\omega_n^2(\theta_0)$ тем, что вместо истинного значения параметра θ_0 подставлена его оценка θ^* .

Распределение статистики $\omega_n^2(\theta^*)$ сильно отличается от распределения классической статистики омега-квадрат $\omega_n^2(\theta_0)$. В качестве примера рассмотрим проверку нормальности, когда $\theta = (m, \sigma^2)$, а $\theta^* = (\bar{x}, s^2)$ (т.е. по выборке оцениваются как математическое ожидание, так и дисперсия нормального распределения). Для этого случая квантили распределений статистик $\omega_n^2(\theta_0)$ и $\omega_n^2(\theta^*)$ приведены в табл.2 (см., например, [19]). Таким образом, квантили отличаются в 3 - 4 раза. Для приведенных в нижней строке табл.2 уровней значимости соответствующие квантили являются критическими значениями при проверке нулевых гипотез, описанных выше.

Таблица 2. Квантили статистики омега-квадрат (Крамера – Мизеса – Смирнова) $\omega_n^2(\theta_0)$ и типа омега-квадрат $\omega_n^2(\theta^*)$ при проверке нормальности

Порядок p квантиля	0,85	0,90	0,95	0,975	0,99
Квантили порядка p для $\omega_n^2(\theta_0)$	0,284	0,347	0,461	0,581	0,743
Квантили порядка p для $\omega_n^2(\theta^*)$	0,091	0,104	0,126	0,148	0,178
Уровень значимости α	0,15	0,10	0,05	0,025	0,01

С сожалением констатируем, что в ряде литературных источников и программных продуктов при описании и использовании критериев согласия допускается грубая ошибка: при проверке нормальности вместо критических значений статистики типа омега-квадрат $\omega_n^2(\theta^*)$ применяют критические значения классической статистики омега-квадрат (Крамера – Мизеса – Смирнова) $\omega_n^2(\theta_0)$. В результате этой грубой ошибки гипотеза нормальности принимается гораздо чаще, чем следует. Эта ошибка подробно рассмотрена в статье [19].

Описанные выше ошибки давно известны специалистам. Ю.В. Линник обращал на них внимание еще в 1953 г. [51, 52]. Суть дела была полностью выяснена в статье Каца, Кифера и Вольфовица [34] 1955 г., содержащей, в частности, материал, представленный в таблицах 1 и 2 выше. (Эту публикацию автор настоящей статьи изучал в 1968 г. качестве курсовой работы на третьем году обучения на кафедре теории вероятностей и математической статистики механико-математического факультета МГУ им. М.В. Ломоносова). В дальнейшем методы расчетов таблиц типа 1 и 2 были развиты, в частности, Ю.Н. Тюриным [38] и его учениками, Г.В. Мартыновым [41]. О рассмотренных выше ошибках Ю.Н. Тюрин рассказал в выпущенной массовым тиражом (40 тыс. экз.) брошюре Общества «Знание» [53]. Наша статья [19] была выпущена в рубрике «Наши консультации» раздела «Математические методы исследования» журнала «Заводская лаборатория» (тираж 9390 экз.).

Однако поток публикаций с ошибками не утихает. Особенно позорным является полностью ошибочный ГОСТ 11.006-74 (СТ СЭВ 1190-78) «Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим», отмененный лишь в 1987 г. – через 13 лет после утверждения, и ставший за это время стандартом СЭВ. Весьма печально, что ошибки встречаются и в учебной литературе. Некоторые примеры приведены в статье [19] и в материалах Интернет-ресурса [54].

Причины, очевидно, лежат в низкой квалификации лиц, берущихся за изготовление учебников. Жаль, что из всех публикаций автора настоящей статьи консультация [19] остается одной из самых востребованных.

6. Заключительные замечания

Критерии Колмогорова, Смирнова, омега-квадрат (Крамера – Мизеса – Смирнова) являются непараметрическими, поскольку при их определении не используются какие-либо параметрические семейства распределений [55].

Распределения статистик этих критериев – одни и те же при любом непрерывном распределении элементов выборки или выборок (при справедливости нулевой гипотезы), т.е. не меняются при любом строго возрастающем непрерывном изменении шкалы измерения. Другими словами, эти критерии предназначены для анализа данных, измеренных в порядковой шкале [50, 56]. Это их свойство вытекает из того, статистики указанных критериев являются функциями не от самых элементов выборки $\xi_1, \xi_2, \dots, \xi_n$, а от величин $F(\xi_1), F(\xi_2), \dots, F(\xi_n)$ где $F(x) = P(\xi_i < x)$ - функция распределения элементов выборки (конкретный вид упомянутых функций см., например, в [18]). Ясно, что

$$P(F(\xi_i) < x) = P(\xi_i < F^{-1}(x)) = F(F^{-1}(x)) = x, \quad 0 \leq x \leq 1,$$

т.е. преобразованные элементы выборки $F(\xi_1), F(\xi_2), \dots, F(\xi_n)$ имеют равномерное распределение на отрезке $[0; 1]$ при любом непрерывном распределении исходных результатов $\xi_1, \xi_2, \dots, \xi_n$ измерений, наблюдений, испытаний, анализов, опытов. Переход от случайной величины ξ к случайной величине $F(\xi)$ называется преобразованием Смирнова. Оно позволяет находить распределения непараметрических критериев с помощью анализа выборок из равномерного распределения на стандартном отрезке $[0; 1]$, а при безграничном росте объема выборки (или

объемов, если выборок несколько) использовать распределения функционалов от таких хорошо изученных случайных процессов, как винеровский процесс и броуновский мост.

Отметим, что значения статистик некоторых критериев типа Колмогорова – Смирнова и типа омега-квадрат меняются при отдельных строго возрастающих преобразованиях шкалы. Таковы, например, критерии типа омега-квадрат, разработанные для проверки симметрии распределения [4, 5]. Однако их распределения (при справедливости нулевой гипотезы) не меняются, поскольку выражаются через распределения выборки из равномерного распределения на отрезке $[0; 1]$, полученной с помощью преобразования Смирнова.

С точки зрения новой парадигмы прикладной математической статистики [57 - 59] и математических методов экономики [60, 61] переход от параметрической статистики Пирсона, Стьюдента и Фишера к непараметрической статистике Колмогорова и Смирнова соответствует переходу от старой парадигмы к новой, от устаревших и не соответствующих реальности постановок задач анализа данных к современным статистическим методам.

Литература

1. Колмогоров А.Н. Метод медиан в теории ошибок // Математический сборник. 1931. Т.38. №3/4. С.47-50.
2. Колмогоров А.Н. К вопросу о пригодности найденных статистическим путем формул прогноза // Журнал геофизики. 1933. Т.3. №1. С.78-82.
3. Смирнов Н.В. О критерии симметрии закона распределения случайной величины // Доклады АН СССР, 1947. Т.56. №1. С.13-16.
4. Орлов А.И. О проверке симметрии распределения // Теория вероятностей и ее применения 1972. Т.17. №2. С.372-377.
5. Орлов А.И. Методы проверки однородности связанных выборок // Заводская лаборатория. Диагностика материалов. 2004. Т.70. №7. С.57-61.
6. Колмогоров А.Н. Теория вероятностей и математическая статистика: Сб. статей. – М.: Наука, 1986. – 535 с.
7. Смирнов Н.В. Теория вероятностей и математическая статистика. Избранные труды. –М.: Наука, 1970. – 289 с.

8. Kolmogoroff A. Sulla determinazione empirica di una legge di distribuzione // *Guornale dell' Instituto Italiano degli Attuari*. 1933. V.4. №1. P.83-91.
9. Колмогоров А.Н. Об одном новом подтверждении законов Менделя // *Доклады АН СССР*, 1940. Т.27. №1. С.38-42.
10. Тутубалин В.Н., Барабашева Ю.М., Девяткова Г.Н., Угер Е.Г. Критерий Колмогорова и экспериментальная проверка законов наследственности Менделя // URL: http://ecology.genebee.msu.ru/3_SOTR/CV_Barabasheva_publ/Kolm-Mend-2008.pdf.
11. Смирнов Н.В. Об отклонениях эмпирической кривой распределения // *Математический сборник*. 1939. Т.6(48). №1. С.3-24.
12. Смирнов Н.В. Оценка расхождения между эмпирическими кривыми распределения в двух независимых выборках // *Бюллетень МГУ*. 1939. Т.2. №2. С.3-14.
13. Орлов А.И. Устойчивость в социально-экономических моделях. — М.: Наука, 1979. — 296 с.
14. Орлов А.И. Организационно-экономическое моделирование : учебник : в 3 ч. Ч. 1. Нечисловая статистика. — М.: Изд-во МГТУ им. Н.Э. Баумана, 2009. — 541 с.
15. Смирнов Н.В. Приближение законов распределения случайных величин по эмпирическим данным // *Успехи математических наук*. 1944. Вып.Х. С.179-206.
16. Смирнов Н.В. О приближении плотностей распределения случайных величин // *Ученые записки МГПИ им. В.П. Потемкина*. 1951. Т. XVI. Вып.3. С.69-96.
17. Смирнов Н.В. Вероятности больших значений непараметрических односторонних критериев согласия // *Труды Математического института АН СССР им. В.А. Стеклова*. 1961. Т.64. С.185-210.
18. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. — М.: ВЦ АН СССР, 1968 (2-е изд.). — 474 с.
19. Орлов А.И. Распространенная ошибка при использовании критериев Колмогорова и омега-квадрат // *Заводская лаборатория*. 1985. Т.51. №1. С.60-62.
20. Бондарев Б. В. О проверке сложных статистических гипотез // *Заводская лаборатория*. 1986. Т. 52. № 10. С. 62–63.
21. Кулинская Е.В., Саввушкина Н. Е. О некоторых ошибках в реализации и применении непараметрических методов в пакете для IBM PC // *Заводская лаборатория*. 1990. Т. 56. № 5. С. 96–99.
22. Орлов А.И. Переход от сумм к интегралам и его применения в изучении асимптотических распределений статистик // *Теория вероятностей и ее применения*. 1973. Т. XVIII. №4. С. 881-883.
23. Холлендер М., Вульф Д. Непараметрические методы статистики / Пер. с англ. Д.С. Шмерлинга; под ред. Ю.П. Адлера и Ю.Н. Тюрина. — М.: Финансы и статистика, 1983. — 518 с.
24. Орлов А.И. О критериях Колмогорова и Смирнова // *Заводская лаборатория*. 1995. Т.61. №7. С.59-61.
25. Cramer H. On the composition of elementary errors // *Skand. Aktuarietidskr.* 1928. V/11/ P/17-34, 141-180.
26. Mises R. von. *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik*. Leipzig; Wien, Deutsche. 1931. S.316-335.
27. Smirnoff N.V. Sur la distribution de ω^2 (criterium de M.R. v. Mises) // *Compt. Rend. Ser. Math. Paris*. 1936. V.202. №6. P.449-452.
28. Смирнов Н.В. О распределении ω^2 - критерия Мизеса // *Математический сборник*. 1937. Т.2 (44). №5. С.973-993.
29. Орлов А.И. Скорость сходимости распределения статистики Мизеса — Смирнова // *Теория вероятностей и ее применения*. 1974. Т.19. №4. С.766-786.

30. Орлов А.И. Асимптотическое поведение статистик интегрального типа // Доклады АН СССР. 1974. Т.219. №4. С. 808-811.
31. Орлов А.И. Асимптотическое поведение статистик интегрального типа // Вероятностные процессы и их приложения. Межвузовский сборник научных трудов. - М.: МИЭМ, 1989. С.118-123.
32. Тюрин Ю.Н. Линейная модель в многомерной непараметрической статистике // Многомерный статистический анализ в социально-экономических исследованиях. - М.: Наука, 1974. С.7-24.
33. Орлов А.И. О критериях согласия с параметрическим семейством // Заводская лаборатория. 1997. Т.63. №5. С. 49-50.
34. Кас М., Kiefer J., Wolfowitz J. On tests of normality and other tests of goodness of fit based on distance methods // Ann. Math. Statist. 1955. V.26. №2. P.189-211.
35. Орлов А.И. О современных проблемах внедрения прикладной статистики и других статистических методов // Заводская лаборатория. 1992. Т.58. № 1. С.67-74.
36. Орлов А.И. Сертификация и статистические методы (обобщающая статья) // Заводская лаборатория. 1997. Т.63. №3. С. 55-62.
37. Залесский Б.А., Ольшевская О.В. О функции распределения статистик омега-квадрат при малых выборках // Заводская лаборатория. 1989. Т.55. №7. С.103-105.
38. Тюрин Ю.Н. Исследования по непараметрической статистике (непараметрические методы и линейная модель). Автореф. дисс. докт. физ.-мат. наук. - М.: МГУ, 1985. - 33 с.
39. Тюрин Ю. Н. О предельном распределении статистик Колмогорова–Смирнова для сложной гипотезы // Известия АН СССР. Сер. «Математика». 1984. Т.48. № 6. С. 1314–1343.
40. Тюрин Ю.Н., Саввушкина Н.Е. Критерии согласия для распределения Вейбулла – Гнеденко // Известия АН СССР. Сер. «Техническая кибернетика». 1984. № 3. С.109-112.
41. Мартынов Г.В. Критерии омега-квадрат. - М.: Наука, 1978. -78 с.
42. Durbin J. Kolmogorov–Smirnov test when parameters are estimated // Lect. Notes Math. 1976. Vol. 566. P. 33–44.
43. Орлов А.И. О нецелесообразности использования итеративных процедур нахождения оценок максимального правдоподобия // Заводская лаборатория. 1986. Т.52. №5. С.67-69.
44. Мирвалиев М., Никулин М.С. Критерии согласия типа хи-квадрат // Заводская лаборатория. 1992. Т.58. № 3. С.52-58.
45. Воинов В.Г. Об оптимальных свойствах критерия Рао-Робсон – Никулина // Заводская лаборатория. 2006 Т.72. №3. С.65-70.
46. Орлов А.И. Метод моментов проверки согласия с параметрическим семейством распределений // Заводская лаборатория. 1989. Т.55. № 10. С.90-93.
47. Рунион Р. Справочник по непараметрической статистике. - М.: Финансы и статистика, 1982. - 198 с.
48. Орлов А.И. Часто ли распределение результатов наблюдений является нормальным? // Заводская лаборатория. 1991. Т.57. №7. С.64-66.
49. Орлов А.И. Статистические пакеты – инструменты исследователя // Заводская лаборатория. 2008. Т.74. №5. С.76-78.
50. Орлов А.И. Прикладная статистика. – М.: Экзамен, 2006. – 671 с.
51. Линник Ю.В. Линейные формы и статистические критерии. 1 – 2 // Украинский математический журнал. 1953. Т.5. №2. С.207-243; №3. С.247-290.

52. Линник Ю.В. Избранные труды. Математическая статистика. – Л.: Наука, 1982. – 284 с.
53. Тюрин Ю.Н. Непараметрические методы статистики. – М.: Знание, 1978. – 64 с.
54. Профессора – невежды готовят себе на смену новых невежд // URL: <http://forum.orlovs.pp.ru/viewtopic.php?f=1&t=548> (дата обращения 01.03.2014).
55. Орлов А.И. Вероятность и прикладная статистика: основные факты: справочник. – М.: КНОРУС, 2010. – 192 с.
56. Луценко Е.В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе «Эйдос» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №08(092). С. 859 – 883. – IDA [article ID]: 0921308058. – Режим доступа: <http://ej.kubagro.ru/2013/08/pdf/58.pdf>
57. Орлов А.И. Новая парадигма прикладной статистики // Статистика и прикладные исследования: сборник трудов Всерос. научн. конф. – Краснодар: Издательство КубГАУ, 2011. – С.206-217.
58. Орлов А.И. Новая парадигма прикладной статистики. - Журнал «Заводская лаборатория. Диагностика материалов». №1, часть I. 2012. Том 78. С.87-93.
59. Орлов А.И. Основные черты новой парадигмы математической статистики / А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №06(090). С.188-214. – IDA [article ID]: 0901306013. – Режим доступа: <http://ej.kubagro.ru/2013/06/pdf/13.pdf>
60. Орлов А.И. Новая парадигма разработки и преподавания организационно-экономического моделирования, эконометрики и статистики в техническом университете // Статистика и прикладные исследования: сборник трудов Всерос. научн. конф. – Краснодар: Издательство КубГАУ, 2011. – С.131-144.
61. Орлов А.И. Новая парадигма математических методов экономики // Экономический анализ: теория и практика. – 2013. – № 36 (339). – С.25–30.

References

1. Kolmogorov A.N. Metod median v teorii oshibok // Matematicheskij sbornik. 1931. T.38. №3/4. S.47-50.
2. Kolmogorov A.N. K voprosu o prigodnosti najdennyh statisticheskim putem formul prognoza // Zhurnal geofiziki. 1933. T.3. №1. S.78-82.
3. Smirnov N.V. O kriterii simmetrii zakona raspredelenija sluchajnoj velichiny // Doklady AN SSSR, 1947. T.56. №1. S.13-16.
4. Orlov A.I. O proverke simmetrii raspredelenija // Teorija verojatnostej i ee primenenija 1972. T.17. №2. S.372-377.
5. Orlov A.I. Metody proverki odnorodnosti svjazannyh vyborok // Zavodskaja laboratorija. Diagnostika materialov. 2004. T.70. №7. S.57-61.
6. Kolmogorov A.N. Teorija verojatnostej i matematicheskaja statistika: Sb. statej. – М.: Nauka, 1986. – 535 s.
7. Smirnov N.V. Teorija verojatnostej i matematicheskaja statistika. Izbrannye trudy. –М.: Nauka, 1970. – 289 s.
8. Kolmogoroff A. Sulla determinazione empirica di una legge di distribuzione // Giornale dell' Istituto Italiano degli Attuari. 1933. V.4. №1. P.83-91.

9. Kolmogorov A.N. Ob odnom novom podtverzhdanii zakonov Mendelja // Doklady AN SSSR, 1940. T.27. №1. S.38-42.
10. Tutubalin V.N., Barabasheva Ju.M., Devjatkova G.N., Uger E.G. Kriterij Kolmogorova i jeksperimental'naja proverka zakonov nasledstvennosti Mendelja // URL: http://ecology.genebee.msu.ru/3_SOTR/CV_Barabasheva_publ/Kolm-Mend-2008.pdf.
11. Smirnov N.V. Ob uklonenijah jempiricheskoy krivoj raspredelenija // Matematicheskij sbornik. 1939. T.6(48). №1. S.3-24.
12. Smirnov N.V. Ocenka rashozhdenija mezhdru jempiricheskimi krivymi raspredelenija v dvuh nezavisimyh vyborkah // Bjulleten' MGU. 1939. T.2. №2. S.3-14.
13. Orlov A.I. Ustojchivost' v social'no-jekonomicheskikh modeljah. — M.: Nauka, 1979. — 296 s.
14. Orlov A.I. Organizacionno-jekonomicheskoe modelirovanie : uchebnik : v 3 ch. Ch. 1. Nechislovaja statistika. — M.: Izd-vo MGTU im. N.Je. Baumana, 2009. — 541 s.
15. Smirnov N.V. Priblizhenie zakonov raspredelenija sluchajnyh velichin po jempiricheskim dannym // Uspehi matematicheskikh nauk. 1944. Vyp.H. S.179-206.
16. Smirnov N.V. O priblizhenii plotnostej raspredelenija sluchajnyh velichin // Uchenye zapiski MGPI im. V.P. Potemkina. 1951. T. XVI. Vyp.3. S.69-96.
17. Smirnov N.V. Verojatnosti bol'shijh znachenij neparametricheskikh odnostoronnih kriteriev soglasija // Trudy Matematicheskogo instituta AN SSSR im. V.A. Steklova. 1961. T.64. S.185-210.
18. Bol'shev L.N., Smirnov N.V. Tablicy matematicheskoy statistiki. — M.: VC AN SSSR, 1968 (2-e izd.). — 474 s.
19. Orlov A.I. Rasprostranennaja oshibka pri ispol'zovanii kriteriev Kolmogorova i omega-kvadrat // Zavodskaja laboratorija. 1985. T.51. №1. S.60-62.
20. Bondarev B. V. O proverke slozhnyh statisticheskikh gipotez // Zavodskaja laboratorija. 1986. T. 52. № 10. S. 62–63.
21. Kulinskaja E.V., Savvushkina N. E. O nekotoryh oshibkah v realizacii i primenenii neparametricheskikh metodov v pakete dlja IBM PC // Zavodskaja laboratorija. 1990. T. 56. № 5. S. 96–99.
22. Orlov A.I. Perehod ot summ k integralam i ego primenenija v izuchenii asimptoticheskikh raspredelenij statistik // Teorija verojatnostej i ee primenenija. 1973. T. XVIII. №4. S. 881-883.
23. Hollender M., Vul'f D. Neparametricheskie metody statistiki / Per. s angl. D.S. Shmerlinga; pod red. Ju.P. Adlera i Ju.N. Tjurina. — M.: Finansy i statistika, 1983. — 518 s.
24. Orlov A.I. O kriterijah Kolmogorova i Smirnova // Zavodskaja laboratorija. 1995. T.61. №7. S.59-61.
25. Cramer H. On the composition of elementary errors // Skand. Aktuarientidskr. 1928. V/11/ P/17-34, 141-180.
26. Mises R. von. Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik. Leipzig; Wien, Deutsche. 1931. S.316-335.
27. Smirnov N.V. Sur la distribution de (criterium de M.R. v. Mises) // Compt. Rend. Ser. Math. Paris. 1936. V.202. №6. P.449-452.
28. Smirnov N.V. O raspredelenii - kriterija Mizesa // Matematicheskij sbornik. 1937. T.2 (44). №5. S.973-993.
29. Orlov A.I. Skorost' shodimosti raspredelenija statistiki Mizesa – Smirnova // Teorija verojatnostej i ee primenenija. 1974. T.19. №4. S.766-786.
30. Orlov A.I. Asimptoticheskoe povedenie statistik integral'nogo tipa // Doklady AN SSSR. 1974. T.219. №4. S. 808-811.

31. Orlov A.I. Asimptoticheskoe povedenie statistik integral'nogo tipa // Veroyatnostnye processy i ih prilozheniya. Mezhvuzovskij sbornik nauchnyh trudov. - M.: MIEM, 1989. S.118-123.
32. Tjurin Ju.N. Linejnaja model' v mnogomernoj neparametricheskoj statistike // Mnogomernyj statisticheskij analiz v social'no-jekonomicheskikh issledovanijah. - M.: Nauka, 1974. S.7-24.
33. Orlov A.I. O kriterijah soglasija s parametricheskim semejstvom // Zavodskaja laboratorija. 1997. T.63. №5. S. 49-50.
34. Kac M., Kiefer J., Wolfowitz J. On tests of normality and other tests of goodness of fit based on distance methods // Ann. Math. Statist. 1955. V.26. №2. P.189-211.
35. Orlov A.I. O sovremennyh problemah vnedrenija prikladnoj statistiki i drugih statisticheskikh metodov // Zavodskaja laboratorija. 1992. T.58. № 1. S.67-74.
36. Orlov A.I. Sertifikacija i statisticheskie metody (obobshhajushhaja stat'ja) // Zavodskaja laboratorija. 1997. T.63. №3. S. 55-62.
37. Zalesskij B.A., Ol'shevskaja O.V. O funkcii raspredelenija statistik omega-kvadrat pri malyh vyborkah // Zavodskaja laboratorija. 1989. T.55. №7. S.103-105.
38. Tjurin Ju.N. Issledovanija po neparametricheskoj statistike (neparametricheskie metody i linejnaja model'). Avtoref. diss. dokt. fiz.-mat. nauk. - M.: MGU, 1985. - 33 s.
39. Tjurin Ju. N. O predel'nom raspredelenii statistik Kolmogorova–Smirnova dlja slozhnoj gipotezy // Izvestija AN SSSR. Ser. «Matematika». 1984. T.48. № 6. C. 1314–1343.
40. Tjurin Ju.N., Savvushkina N.E. Kriterii soglasija dlja raspredelenija Vejbul'la – Gnedenko // Izvestija AN SSSR. Ser. «Tehnicheskaja kibernetika». 1984. № 3. S.109-112.
41. Martynov G.V. Kriterii omega-kvadrat. - M.: Nauka, 1978. -78 s.
42. Durbin J. Kolmogorov–Smirnov test when parameters are estimated // Lect. Notes Math. 1976. Vol. 566. P. 33–44.
43. Orlov A.I. O necelesoobraznosti ispol'zovanija iterativnyh procedur nahozhdenija ocenok maksimal'nogo pravdopodobija // Zavodskaja laboratorija. 1986. T.52. №5. S.67-69.
44. Mirvaliev M., Nikulin M.S. Kriterii soglasija tipa hi-kvadrat // Zavodskaja laboratorija. 1992. T.58. № 3. S.52-58.
45. Voinov V.G. Ob optimal'nyh svojstvah kriterija Rao-Robson – Nikulina // Zavodskaja laboratorija. 2006 T.72. №3. S.65-70.
46. Orlov A.I. Metod momentov proverki soglasija s parametricheskim semejstvom raspredelenij // Zavodskaja laboratorija. 1989. T.55. № 10. S.90-93.
47. Runion R. Spravochnik po neparametricheskoj statistike. - M.: Finansy i statistika, 1982. - 198 s.
48. Orlov A.I. Chasto li raspredelenie rezul'tatov nabljudenij javljaetsja normal'nym? // Zavodskaja laboratorija. 1991. T.57. №7. S.64-66.
49. Orlov A.I. Statisticheskie pakety – instrumenty issledovatelja // Zavodskaja laboratorija. 2008. T.74. №5. S.76-78.
50. Orlov A.I. Prikladnaja statistika. – M.: Jekzamen, 2006. – 671 s.
51. Linnik Ju.V. Linejnye formy i statisticheskie kriterii. 1 – 2 // Ukrainskij matematicheskij zhurnal. 1953. T.5. №2. S.207-243; №3. S.247-290.
52. Linnik Ju.V. Izbrannye trudy. Matematicheskaja statistika. – L.: Nauka, 1982. – 284 s.
53. Tjurin Ju.N. Neparametricheskie metody statistiki. – M.: Znanie, 1978. – 64 s.
54. Professora – nevezhdny gotovjat sebe na smenu novyh nevezhd // URL: <http://forum.orlovs.pp.ru/viewtopic.php?f=1&t=548> (data obrashhenija 01.03.2014).
55. Orlov A.I. Veroyatnost' i prikladnaja statistika: osnovnye fakty: spravochnik. – M.: KNORUS, 2010. – 192 s.

56. Lucenko E.V. Metrizacija izmeritel'nyh shkal razlichnyh tipov i sovmestnaja sopostavimaja kolichestvennaja obrabotka raznorodnyh faktorov v sistemno-kognitivnom analize i sisteme «Jejdos» / E.V. Lucenko // Politematicheskij setевой jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2013. – №08(092). S. 859 – 883. – IDA [article ID]: 0921308058. – Rezhim dostupa: <http://ej.kubagro.ru/2013/08/pdf/58.pdf>

57. Orlov A.I. Novaja paradigma prikladnoj statistiki // Statistika i prikladnye issledovanija: sbornik trudov Vseros. nauchn. konf. – Krasnodar: Izdatel'stvo KubGAU, 2011. – S.206-217.

58. Orlov A.I. Novaja paradigma prikladnoj statistiki. - Zhurnal «Zavodskaja laboratorija. Diagnostika materialov». №1, chast' I. 2012. Tom 78. S.87-93.

59. Orlov A.I. Osnovnye cherty novoj paradigmy matematicheskoy statistiki / A.I. Orlov // Politematicheskij setевой jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2013. – №06(090). S.188-214. – IDA [article ID]: 0901306013. – Rezhim dostupa: <http://ej.kubagro.ru/2013/06/pdf/13.pdf>

60. Orlov A.I. Novaja paradigma razrabotki i prepodavanija organizacionno-jekonomicheskogo modelirovanija, jekonometriki i statistiki v tehničeskom universitete // Statistika i prikladnye issledovanija: sbornik trudov Vseros. nauchn. konf. – Krasnodar: Izdatel'stvo KubGAU, 2011. – S.131-144.

61. Orlov A.I. Novaja paradigma matematicheskikh metodov jekonomiki // Jekonomicheskij analiz: teorija i praktika. – 2013. – № 36 (339). – S.25–30.