УДК 303.732.4 : 519.2

08.00.13 Математические и инструментальные методы экономики (экономические науки)

ЕСТЕСТВЕННЫЕ ПОКАЗАТЕЛИ РАЗЛИЧИЯ

Орлов Александр Иванович
д.э.н., д.т.н., к.ф.-м.н., профессор
РИНЦ SPIN-код: 4342-4994
Московский государственный технический университет им. Н.Э. Баумана, Россия, 105005, Москва, 2-я Бауманская ул., 5, prof-orlov@mail.ru

У статистики нечисловых данных центральная область - статистика в пространствах общей природы (синоним - в пространствах произвольной природы). Для выборок нечисловых данных рассматриваются средние величины и законы больших чисел, статистики интегрального типа, непараметрические оценки плотности, задачи регрессионного и дискриминантного анализа и др. Основополагающие математические понятия - расстояния (метрики, псевдометрики), показатели различия (близости). Аксиоматическое введение расстояний в различных пространствах - популярная тематика на начальном этапе развития статистики нечисловых данных, ведущая начало с книги Кемени - Снелла, в которой аксиоматически введено расстояние между кластеризованными ранжировками. Поэтому целесообразно подробное рассмотрение способов введения метрик и показателей различия, а также изучение их свойств. Попытка выделить свойство, необходимое для получения интересующих нас результатов статистики нечисловых данных, привела нас к понятию "естественного показателя различия" ("естественной меры близости"). Пространства с естественными мерами близости имеют свойства, во многом аналогичные свойствам метрических пространств. В статье приведен ряд соответствующих теорем. Впервые ввёл понятие метрического пространства ввёл Морис Фреше в 1906 г. в связи с рассмотрением функциональных пространств. Неравенство треугольника было унаследовано от евклидовой геометрии. По нашему мнению, нет необходимости в обязательном порядке принимать справедливость неравенства треугольника как аксиому. Другими словами, в различных математических областях целесообразно перейти от метрик (расстояний) к естественным показателям различия. Это в соответствии с новой парадигмой математических методов исследования позволит упростить рассуждения и получать выводы в естественной общности. Актуальным является переход от метрик (расстояний) к естественным показателям различия в различных математических областях. Его можно сопоставить с переходом от классической математики к системной нечетной интервальной математике. Эти два перехода

UDC 303.732.4 : 519.2

08.00.13 Mathematical and instrumental methods of Economics (Economic sciences)

NATURAL INDICATORS OF DIFFERENCE

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

For statistics of non-numerical data, the central area is statistics in spaces of general nature (synonym - in spaces of arbitrary nature). For samples of non-numerical data, averages and laws of large numbers, statistics of integral type, non-parametric density estimates, problems of regression and discriminant analysis, etc. are considered. Fundamental mathematical concepts are distances (metrics, pseudometrics), indicators of difference (proximity). The axiomatic introduction of distances in different spaces is a popular topic at the initial stage of the development of non-numerical data statistics, leading to the Kemeny-Snell book, in which the distance between clustered rankings is axiomatically introduced. Therefore, it is advisable to consider in detail the methods of introducing metrics and indicators of difference, as well as studying their properties. An attempt to isolate the property necessary to obtain the results of statistics of non-numerical data of interest to us, led us to the concept of "natural indicator of difference" ("natural measure of proximity"). Spaces with natural proximity measures have properties that are largely analogous to those of metric spaces. The article contains a number of relevant theorems. Maurice Frechet was the first to introduce the concept of metric space in 1906 in connection with the consideration of functional spaces. The triangle inequality was inherited from Euclidean geometry. In our opinion, there is no need to necessarily accept the validity of the triangle inequality as an axiom. In other words, in various mathematical fields it is advisable to move from metrics (distances) to natural indicators of difference. This, in accordance with the new paradigm of mathematical research methods, will simplify reasoning and draw conclusions in a natural community. The transition from metrics (distances) to natural indicators of difference in various mathematical fields is relevant. It can be compared to the transition from classical mathematics to systematic odd interval mathematics. These two transitions will provide a new dawn of mathematics in the 21st century обеспечат новый рассвет математики в XXI столетии

Ключевые слова: МЕТРИКИ, РАССТОЯНИЯ, МЕРЫ БЛИЗОСТИ, ПОКАЗАТЕЛИ РАЗЛИЧИЯ, АКСИОМЫ, ТОПОЛОГИЯ, КОМПАКТНОСТЬ, НЕПРЕРЫВНОСТЬ, ОПТИМИЗАЦИЯ, ГИСТОГРАММА, ПРЕДЕЛ, ПРИКЛАДНАЯ СТАТИСТИКА

http://dx.doi.org/10.21515/1990-4665-163-020

Keywords: METRICS, DISTANCES, PROXIMITY MEASURES, INDICATORS OF DIFFERENCE, AXIOMS, TOPOLOGY, COMPACTNESS, CONTINUITY, OPTIMIZATION, HISTOGRAM, LIMIT, APPLIED STATISTICS

1. Введение

В настоящее время прикладная статистика [1] продолжает бурно развиваться. Сам термин "прикладная статистика" для обозначения статистических методов анализа данных появился в нашей стране на рубеже 70-80 годов XX в. [2, 3].

Центральная область прикладной статистики - статистика нечисловых данных [4]. Как показано в обзоре [5], среди публикаций 2006-2015 гг. по прикладной статистике 63,0% относится к статистике нечисловых данных.

У статистики нечисловых данных есть своя центральная область - статистика в пространствах общей природы (синоним - в пространствах произвольной природы). Для выборок нечисловых данных, лежащих в некотором пространстве, рассматриваются средние величины и законы больших чисел [6], статистики интегрального типа [7], непараметрические оценки плотности [8 - 10], задачи регрессионного [11] и дискриминантного анализа [12 - 14] и другие постановки.

Во всех этих задачах основополагающие математические понятия - расстояния (метрики, псевдометрики), показатели различия (меры близости). Аксиоматическое введение расстояний в различных пространствах (см., например, [15]) - популярная тематика научных исследований на начальном этапе развития статистики нечисловых данных, ведущая начало со знаменитой книги Кемени - Снелла [16], в которой аксиоматически введено расстояние между кластеризованными ранжировками.

2. Метрики и показатели различия в пространствах нечисловых данных

В статистике нечисловых данных (синонимы - статистике объектов нечисловой природы, нечисловой статистике) постоянно используются функции, измеряющие близость между объектами нечисловой природы [17]. Для их обозначения используют такие термины, как метрика, расстояние, псевдометрика, мера близости, показатель различия. С их помощью определяют средние (теоретические и эмпирические), проводят классификацию (особенно кластер-анализ). Они используются в непараметрических оценках плотности и регрессии, при проверке гипотез о случайных толерантностях и люсианах [18]. В общей схеме устойчивости множество допустимых отклонений и показатели устойчивости также обычно определяют в рассматриваемых терминах [19 - 21]. Сказанное показывает целесообразность подробного рассмотрения способов введения метрик и показателей различия, а также изучения их свойств.

Уточним термины. Пусть X - некоторое множество, $\rho : X^2 \to R^1$.

Определение 1. Функция ρ называется метрикой на X (синоним - расстоянием в X), если для любых x, y, z из X

а) $\rho(x, y) \geq 0$;

б) $\rho(x, y) = \rho(y, x)$;

в) $\rho(x, y) + \rho(y, z) \geq \rho(x, z)$;

г) $\rho(x, x) = 0$;

д) если $\rho(x, y) = 0$, то $x = y$.

Определение 2. Функция ρ называется псевдометрикой на X если для ρ выполнены все перечисленные в определении 1 условия, кроме д).

Иногда псевдометрику также называют метрикой. Это оправдывается возможностью перехода к фактор-пространству $(X \mid \rho)$ путем "склеивания" элементов x и y таких, что $\rho(x, y) = 0$. Однако при рассмотрении нескольких псевдометрик на одном пространстве подобное "склеивание" нецелесообразно.

Определение 3. Функция ρ называется мерой близости на *X* (синоним - показателем различия на *X*), если для ρ выполнены условия а), б) и г) определения 1.

Таким образом, от псевдометрики мера близости отличается отсутствием неравенства треугольника. Так, для случайных величин ξ и η в качестве меры близости можно рассматривать $1 - r_1(\xi, \eta)$ или $1 - r_2(\xi, \eta)$, где $r_1(\xi, \eta)$ - коэффициент линейной парной корреляции Пирсона, $r_2(\xi, \eta)$ - коэффициент ранговой корреляции Спирмена. Неравенство треугольника при этом не выполнено [22]. Этот пример показывает, что меры близости, не являющиеся метриками или псевдометриками, давно использовались в статистике.

Для любой строго возрастающей функции $f : [0, +\infty) \to [0, +\infty)$ такой, что f(0) = 0, функция $f(\rho)$ будет мерой близости, если ρ - мера близости. Переход от ρ к $f(\rho)$ соответствует допустимому преобразованию в порядковой шкале [17]. Адекватность алгоритма обработки данных в порядковой шкале измерения меры близости повышает его обоснованность.

За понятием, введенным в определении 3, в соответствии с рядом литературных источников закрепился термин "мера близости". Он не очень удачен. Термин "близость" иногда порождает лишние ассоциации в социально-экономических приложениях. Более существенно, что чем значение меры близости $\rho(x, y)$ больше, тем близость, похожесть x и y меньше. Поэтому функцию ρ лучше называть "показателем различия.

Иногда [23] мерой близости называют метрику, выведенную из некоторой системы аксиом.

Постулирование справедливости неравенства треугольника - во многом дань традиции. Попытка выделить свойство, необходимое для получения интересующих нас результатов статистики нечисловых данных, привела нас в [24] к понятию "естественного показателя различия" ("естественной меры близости").

Определение 4. Мера близости ρ в пространстве X называется естественной, если из $\rho(x_n, x) \to 0$ и $\rho(y_n, x) \to 0$ вытекает $\rho(x_n, y_n) \to 0$, а из $\rho(x_n, x) \to 0$ и $\rho(x_n, y_n) \to 0$ вытекает $\rho(y_n, x) \to 0$ при $n \to \infty$ для любых $x, x_n, y_n \in X$, n = 1, 2, ...

Теорема 1. Пусть τ - псевдометрика, строго возрастающая функция $f:[0,+\infty) \to [0,+\infty)$ непрерывна в 0, причем $f(0)$ = 0. Тогда $\rho = f(\tau)$ - естественная мера близости.

Доказательство. Рассмотрим множество $A = \{\tau(x, y), y \in X\} \setminus \{0\}$. Либо 0 - предельная точка этого множества, либо нет.

В первом случае существует последовательность $\tau_k \in A, k = 1,2,...$, такая, что $\tau_k \to 0$ при $k \to \infty$, причем $\tau_k = \tau(x, z_k)$. Из непрерывности f следует, что $f(\tau_k) \to 0$ при $k \to \infty$. Без ограничения общности можно считать, что $2\tau_{k+1} < \tau_k$. Пусть $\rho(x_n, x) \to 0$ и $\rho(y_n, x) \to 0$ при $n \to \infty$. Выберем $\varepsilon > 0$. Найдется $k(0)$ такое, что $f(\tau_{k(0)}) < \varepsilon$. Выберем n_0 так, чтобы $\rho(x_n, x) = f(\tau(x_n, x)) < f(\tau_{k(0)+1}), \rho(y_n, x) < f(\tau_{k(0)+1})$ при $n > n_0$. Из строгого возрастания f следует, что $\tau(x_n, x) < \tau_{k(0)+1}, \tau(y_n, x) < \tau_{k(0)+1}$. По неравенству треугольника для τ имеем $\tau(x_n, y_n) \le \tau(x_n, x) + \tau(x, y_n) < 2\tau_{k(0)+1} < \tau_{k(0)}$. Из строгого возрастания f вытекает, что $\rho(x_n, y_n) = f(\tau(x_n, y_n)) < f(\tau_{k(0)}) < \varepsilon$ при $n > n_0$. Следовательно, $\rho(x_n, y_n) \to 0$ при $n \to \infty$. Второе соотношение в определении 4 доказывается аналогично.

Если же A отделено от 0, то $X = X_1 \cup X_2$, где $\tau(x,y)=0$ при $y \in X_1$ и $\tau(x,y) \ge \delta_0 > 0$ при $y \in X_2$. Из строгого возрастания f следует, что $\rho(x,y) \ge f(\delta_0)$ при $y \in X_2$. Следовательно, из $\rho(x_n,x) \to 0$ и $\rho(y_n,x) \to 0$ вытекает, что существует n_0 такое, что $x_n \in X_1$ и $y_n \in X_1$ при $n > n_0$. Тогда $\tau(x_n,y_n) \le \tau(x_n,x)+\tau(x,y_n)=0$, что и требовалось доказать. Второе соотношение определении 4 требует специального рассмотрения. Как и раньше, из $\rho(x_n,x) \to 0$ вытекает, что $x_n \in X_1$ при $n > n_0$. Из $\rho(x_n,y_n) \to 0$ с учетом строгой монотонности f следует, что $\tau(x_n,x) \to 0$, значит, $\tau(x_n,y_n) < \delta_0$ при $n > n_1$. Тогда по неравенству треугольника $\tau(y_n,x) \le \tau(x_n,y_n)+\tau(x_n,x) < \delta_0$, а потому $\tau(y_n,x)=0$, откуда вытекает $y_n \in X_1$ и $f(\tau(y_n,x))=0$. Теорема 1 доказана.

Близкие вопросы в задачах кластер-анализа рассмотрены в [25]. Таким образом, понятие "естественной меры близости" введено непосредственно из прикладных соображений (работа [25] написана нами, теорема 1 и приведенные ниже результаты получены нами).

3. Показатели различия в прикладной статистике

В прикладной статистике используют весьма большое число различных показателей различия, метрик и мер близости (см. сводки в [26, 27]). Возникает вопрос: какой из них в конкретной задаче пользоваться?

В 1959 г. Дж. Кемени предложил аксиоматический подход, согласно которому следует сформулировать естественные для конкретной задачи аксиомы и вывести из них вид метрики [28]. Этот подход получил большую популярность в нашей стране после выхода в 1972 г. перевода на русский язык книги Дж. Кемени и Дж. Снелла [16], в которой дана система аксиом для расстояния Кемени между упорядочениями. Отметим, что определения терминов "расстояние Кемени" и "медиана Кемени" даны в наших одноименных статьях [29, 30] в Энциклопедии "Вероятность и математическая статистика".

Последовала большая серия работ, в которых из более или менее аналогичных систем аксиом выводились метрики или псевдометрики для различных объектов нечисловой природы. В частности, нами были предложены системы аксиом и выведены расстояния в пространствах толерантностей и множеств, совместно с Г.В. Раушенбахом - в пространствах неотрицательных суммируемых функций с мерами [15, 17]. Сводка результатов, полученных при аксиоматическом введении метрик, дана в обзоре [31]. Он содержит 161 литературную ссылку, в том числе 69 ссылок на русском языке. Этот обзор дает представление о развитии исследований по аксиоматическому введению метрик (хотя в нем и отсутствуют некоторые важные ссылки на литературные источники, например, на [32] или [33]), что избавляет нас от необходимости давать здесь еще один обзор по рассматриваемой тематике. Отметим только, что первоначальная идея Дж. Кемени в известной мере изжила себя: раньше мы имели много метрик и надеялись с помощью аксиоматического подхода выбрать из них наиболее подходящую, теперь же наряду с конкретной метрикой мы в типичных случаях имеем и систему аксиом, ее порождающую, и вместо выбора метрики вынуждены выбирать среди систем аксиом. Как уже отмечалось, мы также отдали дань "моде" на аксиоматическое введение метрик - три системы аксиом, порождающие метрики и псевдометрики в пространствах толерантностей, множеств и неотрицательных суммируемых функций соответственно, приведены в [15, 17] и других наших публикациях.

Ряд свойств метрик переносится на естественные показатели различия (меры близости). Рассмотрим их.

4. Естественные показатели различия

Пространства с естественными мерами близости, они же - пространства с естественными показателями различия, имеют свойства, во многом аналогичные свойствам метрических пространств. Так, известно, что всякое счетно-компактное метрическое пространство бикомпактно (см., например, [34, с.102]. Аналог этого утверждения доказан в [35].

Теорема 2. Пусть X счетно-компактно, топология порождена естественной мерой близости ρ. Тогда X бикомпактно.

Доказательство. Понадобится понятие ε-сети. Для подмножества естественной мерой близости) ε-сеть - это множество из того же пространства такое, что для любой точки найдётся точка удалённая от не более чем на ε.

Покажем сначала, что в X = для любого ε > 0 существует конечная ε-сеть. Предположим противное - для какого-то $\varepsilon_0 > 0$ её нет. Тогда мы можем построить бесконечную последовательность элементов a_n из X, n = 1, 2, ..., такую, что

$$\rho(a_i, a_j) \geq \varepsilon_0, i \neq j, i, j = 1,2,... \quad (1)$$

В силу счетной компактности X у последовательности $\{a_n\}$ существует предельная точка x, т.е. для некоторой последовательности натуральных чисел $n(k)$, k = 1, 2, ..., имеем

$$\lim_{k\to\infty} a_{n(k)} = x. \quad (2)$$

Положим

$$x_k = a_{n(2k)},\ y_k = a_{n(2k+1)},\ k = 1, 2, ... \quad (3)$$

Тогда в силу (2) и (1) имеем

$$\rho(x_k, y_k) \geq \varepsilon_0,\ \ \rho(x_k, x) \to 0,\ \ \rho(y_k, x) \to 0, \quad (4)$$

что противоречит определению естественной меры близости. Следовательно, при любом $\varepsilon > 0$ существует конечная ε-сеть.

Пространство X является сепарабельным: счетным всюду плотным множеством является объединение всех $1/n$ - сетей, n = 1, 2, ... Покажем, что X имеет счетную базу. Рассмотрим совокупность шаров радиуса $1/n$ с центрами в x_m, где m, n = 1, 2, ..., а $\{x_m, m = 1, 2, ...\}$ - счетное всюду плотное множество. Покажем, что эта совокупность есть база. Для этого в соответствии с теоремой 3 [34, с.83] достаточно показать, что каждого открытого множества G и каждой точки $x \in G$ существует шар G_x из указанной совокупности такой, что $x \in G_x \subset G$. Действительно, это так: в силу того, что топология порождена ρ, существует $\varepsilon > 0$ такое, что шар G_1 радиуса ε с центром в x полностью лежит в G.

Рассмотрим последовательность точек из счетного всюду плотного множества, сходящуюся к x, и последовательность шаров с центрами в этих точках и минимальными радиусами вида $1/m$, m = 1, 2, ..., содержащих x (т.е. для шара с центром в y имеем $m = \max\{k : 1/k > \rho(x, y)\}$). Предположим, что ни один из них не лежит в G_1. Пусть x_n - центры шаров, α_n - их радиусы, y_n - та точка n-го шара, что лежит вне G_1. Тогда

$$\alpha_n > \rho(x_n, x) \to 0,\ \alpha_n > \rho(x_n, y_n) \to 0,\ \rho(x, y_n) \geq \varepsilon, \tag{5}$$

что противоречит определению естественной меры близости.

Таким образом, счетная база существует. Для завершения доказательства теоремы 2 достаточно сослаться на теорему 10 из [34, с.98].

Определение 5. Пусть X и Y - пространства с мерами близости ρ_1 и ρ_2 соответственно. Функция $f : X \to Y$ называется равномерно непрерывной в X в смысле мер близости ρ_1 и ρ_2, если для любого $\varepsilon > 0$ найдется $\delta > 0$ такое, что для любых x, x' таких, что $\rho_1(x, x') < \delta$, имеем $\rho_2(f(x), f(x')) < \varepsilon$.

Известно, что непрерывное отображение метрического компакта в метрическое пространство равномерно непрерывно (теорема 6 в [34, с.108]). Справедлив аналог этого утверждения.

Теорема 3. Пусть X счетно-компактно, топология в X порождена естественной мерой близости ρ_1, в пространстве Y задана естественная мера близости ρ_2, функция $f: X \to Y$ непрерывна. Тогда f равномерно непрерывна.

Доказательство. Предположим, что отображение f не является равномерно непрерывным. Это означает, что для некоторого $\varepsilon > 0$ и каждого натурального n найдутся в X такие точки x_n и x_n', что

$$\rho_1(x_n, x_n') < 1/n, \ \rho_2(f(x_n), f(x_n')) \geq \varepsilon. \qquad (6)$$

Бесконечное подмножество $\{x_n \ , \ n = 1, 2, ...\}$ имеет хотя бы одну предельную точку. Значит, существуют $x_0 \in X$ и последовательность $n(k)$, $k = 1, 2, ...$ такие, что $x_{n(k)} \to x_0$ при $k \to \infty$. Тогда в силу первого неравенства в (6) и определения естественной меры близости справедливо предельное соотношение $x'_{n(k)} \to x_0$. Из определения непрерывности f следует, что

$$\rho_2(f(x_{n(k)}), f(x_0)) \to 0, \rho_2(f(x'_{n(k)}), f(x_0)) \to 0, \qquad (7)$$

а это в силу определения естественной меры близости несовместимо со вторым неравенством в (6). Теорема 3 доказана.

Дальнейшие рассмотрения непосредственно ориентированы на изучение естественных показателей различия (естественных мер близости) как составляющих непараметрических оценок плотности [1, 8, 9, 17].

Определение 6. Разбиением $R = R(X)$ пространства X называется конечная совокупность подмножеств $R = \{X_1, X_2, ..., X_k\}$, называемых областями, или элементам разбиения, такая, что

$$X_1 \cup X_2 \cup ... \cup X_k = X, \ X_i \cap X_j = \varnothing, \ i \neq j, \ i, j = 1,2,...k.$$

Определение 7. Гистограммой $f^*(x)$ функции $f(x)$ называется функция $f^*: X \to R^1$ такая, что

$$f^*(x) = f(x_m), \ x \in X_m, \ m = 1,2,...,k.$$

Таким образом, для задания гистограммы необходимо указать разбиение $R = \{X_1, X_2, ..., X_k\}$ и набор точек $x_1, x_2, ..., x_k$, по одной из каждого элемента разбиения, $x_i \in X_i$, $i = 1, 2, ..., k$.

Будем изучать условия, при которых

$$\sup_{x \in X} | f(x) - f^*(x) | \to 0 \quad (8)$$

при измельчении разбиений. Под последовательностью измельчающихся разбиений будем понимать такую последовательность, в которой каждое следующее разбиение получается из предыдущего разбиением на части некоторых из входящих в него областей. Ясно, что, зная f, легко построить измельчающуюся последовательность, для которой справедливо (8). Однако для нас важен случай, когда f неизвестна, ее необходимо оценить. Полезным окажется следующее условие.

Условие А. Существует последовательность разбиений, измельчающаяся и такая, что для любой точки $x \in X$ и любой ее окрестности $U(x)$ найдется такое разбиение из этой последовательности, что его область, содержащая x, полностью содержится в $U(x)$.

Теорема 4. Пусть f - непрерывная функция, X счетно-компактно, выполнено условие А. Тогда справедливо (8).

Доказательство. Предположим, что (8) неверно. Это означает, что для некоторого $\varepsilon_0 > 0$ при любом $n = 1, 2, ...$ существует элемент $x_n \in X$ такой, что

$$| f(x_n) - f^*(x_n) | > \varepsilon_0, \quad (9)$$

где f^* строится по $m(n)$-му разбиению описанной в условии А последовательности. В силу счетной компактности X у последовательности $\{x_n, n = 1, 2, ...\}$ существует предельная точка x_0, т.е. $x_{n(k)} \to x_0$ для некоторой последовательности $n(k)$, $k = 1, 2, ...$ Выберем окрестность $U(x_0)$ такую, что $| f(y) - f(x_0) | < \varepsilon_0/2$ при $y \in U(x_0)$. В силу условия

А найдется область одного (а потому и всех последующих) из разбиений, полностью лежащая в $U(x_0)$. Пусть это разбиение имеет номер n_0. Тогда для всех k таких, что $n(k) > n_0$, значение $f^*(x_{n(k)})$ определяется по области, полностью лежащей в $U(x_0)$, а потому

$$\lim_{k\to\infty} f^*(x_{n(k)}) = f(x_0),\ \lim_{k\to\infty} |f(x_{n(k)}) - f^*(x_{n(k)})| = 0, \tag{10}$$

что противоречит (9). Теорема 4 доказана.

Замечание. Если топология в X определяется с помощью естественной меры близости, то согласно теореме 3 функция $f: X \to Y$ равномерно непрерывна, и теорема 4 вытекает из теоремы 3.

Теорема 5. Пусть X счетно-компактно, топология порождена естественной мерой близости ρ. Тогда условие А выполнено.

Доказательство. Как показано при доказательстве теоремы 2, для любого n = 1, 2, ... в X существует конечная $1/n$-сеть. Возьмем шары радиуса $1/n$ с центрами в точках $1/n$-сети. Их объединение совпадает с X. Все их непустые пересечения составляют в совокупности разбиение X. Чтобы получить последовательность измельчающихся разбиений, достаточно рассмотреть в качестве n-го члена последовательности совокупность всех непустых пересечений областей описанного в предыдущей фразе разбиения, построенного по $1/n$-сети, с областями аналогичных разбиений, построенных по $1/m$-сетям, $m < n$ (аналогичное построение подробно проведено в [19, с.272-273]).

Возьмем $x \in X$ и его окрестность $U(x)$. Поскольку объединение всех $1/n$-сетей, n = 1, 2, ... , есть всюду плотное множество, то из этого множества можно выбрать последовательность x_k, k = 1, 2, ... , сходящуюся к x при $k \to \infty$. Пусть S_k - область разбиения, соответствующая x_k как элементу $1/m$-сети при некотором m. Покажем, что $S_k \subseteq U(x)$ при некотором k. Поскольку топология в X порождена мерой близости ρ, то существует $\varepsilon > 0$ такое, что шар радиуса ε с центром в x входит в $U(x)$. Предположим, что $S_k \setminus U(x) \neq \varnothing$ при всех k. Это означает, что существует последовательность y_k, k = 1, 2, ... , $y_k \in S_k$, такая, что $\rho(y_k, x) \geq \varepsilon$ при всех k. Вместе с тем поскольку $S_k \subseteq \{x : \rho(x, x_k) \leq 1/m\}$ и $m \to \infty$ при $k \to \infty$, то $\rho(x_k, y_k) \to 0$ (и $\rho(x_k, x) \to 0$ по построению x_k) при $k \to \infty$. Получили противоречие с определением естественной меры близости. Теорема 5 доказана.

5. Заключительные замечания

Метрические пространства, различные их вариации и обобщения - важная составляющая инструментария современной математики. В частности, в статистике нечисловых данных метрики (расстояния), показатели различия и меры близости используются для определения средних величин и доказательства законов больших чисел [6, 17], изучения статистик интегрального типа [7], формирования непараметрические оценки плотностей в пространствах нечисловой природы [8 - 10]. На них основаны постановки и методы решения задач восстановления зависимости (регрессионного анализа), как параметрических, так и непараметрических [11]. В задачах классификации задачи кластер-анализа формулируются в терминах метрик (расстояний), показателей различия и мер близости, они необходимы в диагностике и дискриминантном анализе [12 - 14], во многих других постановках.

В ходе исторического развития математики метрические пространства появились в начале XX в. По-видимому, впервые ввёл понятие метрического пространства Морис Фреше [36] в связи с рассмотрением функциональных пространств. Неравенство треугольника было унаследовано от евклидовой геометрии. По нашему мнению, нет необходимости в обязательном порядке принимать справедливость неравенства треугольника как аксиому. Другими словами, в различных математических областях целесообразно перейти от метрик (расстояний) к естественным показателям различия. Это в соответствии с новой парадигмой математических методов исследования [37] позволит упростить рассуждения и получать выводы в естественной общности.

В настоящее время актуальной проблемой является переход от метрик (расстояний) к естественным показателям различия в различных математических областях. Его можно сопоставить с переходом от классической математики к системной нечетной интервальной математике, намеченным в монографии [38]. Эти два перехода обеспечат новый рассвет математики в XXI столетии.

Литература

1. Орлов А.И. Прикладная статистика. — М.: Экзамен, 2006. — 671 с.
2. Орлов А.И. Непараметрическая и прикладная статистика в нашей стране / Научный журнал КубГАУ. 2014. №101. С. 197–226.
3. Лойко В.И., Луценко Е.В., Орлов А.И. Высокие статистические технологии и системно-когнитивное моделирование в экологии. – Краснодар : КубГАУ, 2019. – 258 с.
4. Орлов А.И. Статистика нечисловых данных - центральная часть современной прикладной статистики / Научный журнал КубГАУ. 2020. № 156. С. 111–142.
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. Орлов А.И. Средние величины и законы больших чисел в пространствах произвольной природы / Научный журнал КубГАУ. 2013. №89. С. 556–586.
7. Орлов А.И. Предельная теория непараметрических статистик / Научный журнал КубГАУ. 2014. №100. С. 224–242.
8. Орлов А.И. Оценки плотности распределения вероятностей в пространствах произвольной природы / Научный журнал КубГАУ. 2014. №99. С. 33–49.
9. Орлов А.И. Предельные теоремы для ядерных оценок плотности в пространствах произвольной природы / Научный журнал КубГАУ. 2015. №108. С. 316–333.
10. Орлов А.И. Асимптотика оценок плотности распределения вероятностей / Научный журнал КубГАУ. 2017. №131. С. 832–860.
11. Орлов А.И. Предельная теория решений экстремальных статистических задач / Научный журнал КубГАУ. 2017. №133. С. 579–600..
12. Орлов А.И. Математические методы теории классификации / Научный журнал КубГАУ. 2014. №95. С. 423–459.
13. Орлов А.И. Базовые результаты математической теории классификации / Научный журнал КубГАУ. 2015. №110. С. 219–239.
14. Орлов А.И. Прогностическая сила – наилучший показатель качества алгоритма диагностики / Научный журнал КубГАУ. 2014. №99. С. 15–32.

15. Орлов А.И. Расстояния в пространствах статистических данных / Научный журнал КубГАУ. 2014. №101. С. 227–252.
16. Кемени Дж., Снелл Дж. Кибернетическое моделирование. Некоторые приложения. - М.: Советское радио, 1972. - 192 с.
17. Орлов А.И. Организационно-экономическое моделирование: : учебник : в 3 ч. Ч.1: Нечисловая статистика. — М.: Изд-во МГТУ им. Н. Э. Баумана, 2009. — 542 с.
18. Орлов А.И. Теория люсианов / Научный журнал КубГАУ. 2014. №101. С. 275–304.
19. Орлов А.И. Устойчивость в социально-экономических моделях. — М.: Наука, 1979. — 296 с.
20. Орлов А.И. Устойчивые экономико-математические методы и модели. Разработка и развитие устойчивых экономико-математических методов и моделей для модернизации управления предприятиями. — Saarbrücken (Germany), LAP (Lambert Academic Publishing), 2011. — 436 с.
21. Орлов А.И. Новый подход к изучению устойчивости выводов в математических моделях / Научный журнал КубГАУ. 2014. №100. С. 1 – 30.
22. Кузьмин В.Б., Овчинников С.В. Модель для измерений в порядковых шкалах / Многомерный статистический анализ в социально-экономических исследованиях. - М.: Наука, 1974. - С.384-388.
23. Литвак Б.Г. Экспертная информация: Методы получения и анализа. - М.: Радио и связь, 1982. - 184 с.
24. Орлов А.И. Непараметрические оценки плотности в топологических пространствах / Прикладная статистика. - М.: Наука, 1983. - С. 12-40.
25. Толстова Ю.Н. Адекватность функции расстояния в алгоритмах автоматической классификации // Исследования по вероятностно-статистическому моделированию реальных систем. - М.: ЦЭМИ АН СССР, 1977. - С. 168-173.
26. Воронин Ю.А. Теория классифицирования и ее приложения. - Новосибирск: Наука, 1985. - 232 с.
27. Sokal R.R., Sneath P.H.A. Principles of numerical taxonomy. - San Francisco - London, W.H. Freeman and company, 1963. - 359 pp.
28. Kemeny J. Mathematics without numbers / Daedalus. 1959. V.99. P. 571-591.
29. Орлов А.И. Кемени расстояние // Вероятность и математическая статистика. Энциклопедия / Гл. ред. Ю. В. Прохоров. – М.: Изд-во «Большая Российская Энциклопедия», 1999. - С. 230.
30. Орлов А.И. Кемени медиана // Вероятность и математическая статистика. Энциклопедия / Гл. ред. Ю. В. Прохоров. – М.: Изд-во «Большая Российская Энциклопедия», 1999. - С. 229-230.
31. Раушенбах Г.В. Меры близости и сходства / Анализ нечисловой информации в социологических исследованиях. - М.: Наука, 1985. - С. 169-203.
32. Кузьмин В.Б., Овчинников С.В. Модель для измерений в порядковых шкалах / Многомерный статистический анализ в социально-экономических исследованиях. - М.: Наука, 1974. - С. 384-388.
33. Куликов С.М. Структурные меры близости в пространствах классификаций и разбиений / Прикладная статистика. - М.: Наука, 1983. - С.282-286.
34. Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. - М.: Наука, 1972. - 496 с.
35. Орлов А.И. Непараметрические оценки плотности в топологических пространствах / Прикладная статистика. - М.: Наука, 1983. - С. 12-40.
36. Fréchet M. Sur quelques points du calcul fonctionnel / Rendiconti del Circolo Matematico di Palermo. 1906. V. 22. P. 1—74.

37. Орлов А.И. О новой парадигме математических методов исследования / Научный журнал КубГАУ. 2016. №122. С. 807–832.
38. Орлов А.И., Луценко Е.В. Системная нечеткая интервальная математика. – Краснодар, КубГАУ. 2014. – 600 с.

References

1. Orlov A.I. Prikladnaya statistika. — M.: Ekzamen, 2006. — 671 s.
2. Orlov A.I. Neparametricheskaya i prikladnaya statistika v nashej strane / Nauchnyj zhurnal KubGAU. 2014. №101. S. 197–226.
3. Lojko V.I., Lucenko E.V., Orlov A.I. Vysokie statisticheskie tekhnologii i sistemno-kognitivnoe modelirovanie v ekologii. – Krasnodar : KubGAU, 2019. – 258 s.
4. Orlov A.I. Statistika nechislovyh dannyh - central'naya chast' sovremennoj prikladnoj statistiki / Nauchnyj zhurnal KubGAU. 2020. № 156. S. 111–142.
5. Orlov A.I. Razvitie matematicheskih metodov issledovaniya (2006 – 2015 gg.) // Zavodskaya laboratoriya. Diagnostika materialov. 2017. T.83. №1. CH.1. S. 78-86.
6. Orlov A.I. Srednie velichiny i zakony bol'shih chisel v prostranstvah proizvol'noj prirody / Nauchnyj zhurnal KubGAU. 2013. №89. S. 556–586.
7. Orlov A.I. Predel'naya teoriya neparametricheskih statistik / Nauchnyj zhurnal KubGAU. 2014. №100. S. 224–242.
8. Orlov A.I. Ocenki plotnosti raspredeleniya veroyatnostej v prostranstvah proizvol'noj prirody / Nauchnyj zhurnal KubGAU. 2014. №99. S. 33–49.
9. Orlov A.I. Predel'nye teoremy dlya yadernyh ocenok plotnosti v prostranstvah proizvol'noj prirody / Nauchnyj zhurnal KubGAU. 2015. №108. S. 316–333.
10. Orlov A.I. Asimptotika ocenok plotnosti raspredeleniya veroyatnostej / Nauchnyj zhurnal KubGAU. 2017. №131. S. 832–860.
11. Orlov A.I. Predel'naya teoriya reshenij ekstremal'nyh statisticheskih zadach / Nauchnyj zhurnal KubGAU. 2017. №133. S. 579–600..
12. Orlov A.I. Matematicheskie metody teorii klassifikacii / Nauchnyj zhurnal KubGAU. 2014. №95. S. 423–459.
13. Orlov A.I. Bazovye rezul'taty matematicheskoj teorii klassifikacii / Nauchnyj zhurnal KubGAU. 2015. №110. S. 219–239.
14. Orlov A.I. Prognosticheskaya sila – nailuchshij pokazatel' kachestva algoritma diagnostiki / Nauchnyj zhurnal KubGAU. 2014. №99. S. 15–32.
15. Orlov A.I. Rasstoyaniya v prostranstvah statisticheskih dannyh / Nauchnyj zhurnal KubGAU. 2014. №101. S. 227–252.
16. Kemeni Dzh., Snell Dzh. Kiberneticheskoe modelirovanie. Nekotorye prilozheniya. - M.: Sovetskoe radio, 1972. - 192 s.
17. Orlov A.I. Organizacionno-ekonomicheskoe modelirovanie: : uchebnik : v 3 ch. CH.1: Nechislovaya statistika. — M.: Izd-vo MGTU im. N. E. Baumana, 2009. — 542 s.
18. Orlov A.I. Teoriya lyusianov / Nauchnyj zhurnal KubGAU. 2014. №101. S. 275–304.
19. Orlov A.I. Ustojchivost' v social'no-ekonomicheskih modelyah. — M.: Nauka, 1979. — 296 s.
20. Orlov A.I. Ustojchivye ekonomiko-matematicheskie metody i modeli. Razrabotka i razvitie ustojchivyh ekonomiko-matematicheskih metodov i modelej dlya modernizacii upravleniya predpriyatiyami. — Saarbrücken (Germany), LAP (Lambert Academic Publishing), 2011. — 436 s.
21. Orlov A.I. Novyj podhod k izucheniyu ustojchivosti vyvodov v matematicheskih modelyah / Nauchnyj zhurnal KubGAU. 2014. №100. S. 1 – 30.

22. Kuz'min V.B., Ovchinnikov S.V. Model' dlya izmerenij v poryadkovyh shkalah / Mnogomernyj statisticheskij analiz v social'no-ekonomicheskih issledovaniyah. - M.: Nauka, 1974. - S.384-388.
23. Litvak B.G. Ekspertnaya informaciya: Metody polucheniya i analiza. - M.: Radio i svyaz', 1982. - 184 s.
24. Orlov A.I. Neparametricheskie ocenki plotnosti v topologicheskih prostranstvah / Prikladnaya statistika. - M.: Nauka, 1983. - S. 12-40.
25. Tolstova YU.N. Adekvatnost' funkcii rasstoyaniya v algoritmah avtomaticheskoj klassifikacii // Issledovaniya po veroyatnostno-statisticheskomu modelirovaniyu real'nyh sistem. - M.: CEMI AN SSSR, 1977. - S. 168-173.
26. Voronin YU.A. Teoriya klassificirovaniya i ee prilozheniya. - Novosibirsk: Nauka, 1985. - 232 s.
27. Sokal R.R., Sneath P.H.A. Principles of numerical taxonomy. - San Francisco - London, W.H. Freeman and company, 1963. - 359 pp.
28. Kemeny J. Mathematics without numbers / Daedalus. 1959. V.99. P. 571-591.
29. Orlov A.I. Kemeni rasstoyanie // Veroyatnost' i matematicheskaya statistika. Enciklopediya / Gl. red. YU. V. Prohorov. – M.: Izd-vo «Bol'shaya Rossijskaya Enciklopediya», 1999. - S. 230.
30. Orlov A.I. Kemeni mediana // Veroyatnost' i matematicheskaya statistika. Enciklopediya / Gl. red. YU. V. Prohorov. – M.: Izd-vo «Bol'shaya Rossijskaya Enciklopediya», 1999. - S. 229-230.
31. Raushenbah G.V. Mery blizosti i skhodstva / Analiz nechislovoj informacii v sociologicheskih issledovaniyah. - M.: Nauka, 1985. - S. 169-203.
32. Kuz'min V.B., Ovchinnikov S.V. Model' dlya izmerenij v poryadkovyh shkalah / Mnogomernyj statisticheskij analiz v social'no-ekonomicheskih issledovaniyah. - M.: Nauka, 1974. - S. 384-388.
33. Kulikov S.M. Strukturnye mery blizosti v prostranstvah klassifikacij i razbienij / Prikladnaya statistika. - M.: Nauka, 1983. - S.282-286.
34. Kolmogorov A.N., Fomin S.V. Elementy teorii funkcij i funkcional'nogo analiza. - M.: Nauka, 1972. - 496 s.
35. Orlov A.I. Neparametricheskie ocenki plotnosti v topologicheskih prostranstvah / Prikladnaya statistika. - M.: Nauka, 1983. - S. 12-40.
36. Fréchet M. Sur quelques points du calcul fonctionnel / Rendiconti del Circolo Matematico di Palermo. 1906. V. 22. P. 1—74.
37. Orlov A.I. O novoj paradigme matematicheskih metodov issledovaniya / Nauchnyj zhurnal KubGAU. 2016. №122. S. 807–832.
38. Orlov A.I., Lucenko E.V. Sistemnaya nechetkaya interval'naya matematika. – Krasnodar, KubGAU. 2014. – 600 s.