УДК 519.2:303.732.4

UDC 519.2:303.732.4

01.00.00 Физико-математические науки

ВЕРОЯТНОСТНЫЕ МОДЕЛИ ПОРОЖДЕНИЯ НЕЧИСЛОВЫХ ДАННЫХ

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

Статистика объектов нечисловой природы (статистика нечисловых данных, нечисловая статистика) является областью математической статистики, посвященной методам анализа нечисловых данных. Основой применения результатов математической статистики являются вероятностно-статистические модели реальных явлений и процессов, важнейшей (а часто и единственной) составной частью которых являются модели порождения данных. Простейшим примером модели порождения данных является модель выборки как совокупности независимых одинаково распределенных случайных величин. В настоящей статье рассмотрены основные вероятностные модели порождения нечисловых данных. А именно, модели дихотомических данных, результатов парных сравнений, бинарных отношений, рангов, объектов общей природы. Обсуждаются различные варианты вероятностных моделей и их практическое использование. Например, базовая вероятностная модель дихотомических данных - бернуллиевский вектор (люсиан), т.е. конечная последовательность независимых испытаний Бернулли, для которых вероятности успеха могут быть различны. Математический аппарат решения различных статистических задач, связанных с бернуллиевскими векторами, полезен для анализа случайных толерантностей; случайных множества с независимыми элементами; при обработке результатов независимых парных сравнений; в статистических методах анализа точности и стабильности технологических процессов; при анализе и синтезе планов статистического приемочного контроля (по дихотомическим признакам); при обработке маркетинговых и социологических анкет (с закрытыми вопросами типа «да» - «нет»); при обработке социально-психологических и медицинских данных, в частности, ответов на психологические тесты типа MMPI (используемых, в частности, в задачах управления персоналом), при анализе топографических карт (применяемых для анализа и прогноза зон поражения при технологических авариях, распространении коррозии,

PROBABILITY MODELS FOR OBTAINING NON-NUMERICAL DATA

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

The statistics of objects of non-numerical nature (statistics of non-numerical objects, non-numerical data statistics, non-numeric statistics) is the area of mathematical statistics, devoted to the analysis methods of non-numeric data. Basis of applying the results of mathematical statistics are probabilistic-statistical models of real phenomena and processes, the most important (and often only) which are models for obtaining data. The simplest example of a model for obtaining data is the model of the sample as a set of independent identically distributed random variables. In this article we have considered the basic probabilistic models for obtaining non-numeric data. Namely, the models of dichotomous data, results of paired comparisons, binary relations, ranks, the objects of general nature. We have discussed the various options of probabilistic models and their practical use. For example, the basic probabilistic model of dichotomous data - Bernoulli vector (Lucian) i.e. final sequence of independent Bernoulli trials, for which the probabilities of success may be different. The mathematical tools of solutions of various statistical problems associated with the Bernoulli vectors are useful for the analysis of random tolerances; random sets with independent elements; in processing the results of independent pairwise comparisons; statistical methods for analyzing the accuracy and stability of technological processes; in the analysis and synthesis of statistical quality control plans (for dichotomous characteristics); the processing of marketing and sociological questionnaires (with closed questions like "yes" - "no"); the processing of socio-psychological and medical data, in particular, the responses to psychological tests such as MMPI (used in particular in the problems of human resource management), and analysis of topographic maps (used for the analysis and prediction of the affected areas for technological disasters, distributing corrosion, propagation environmentally harmful pollutants, various diseases (including myocardial infarction), in other situations), etc.

распространении экологически вредных загрязнений, различных заболеваниях (в частности, при инфаркте миокарда), в других ситуациях), и т.д.

Ключевые слова: МАТЕМАТИЧЕСКАЯ СТАТИСТИКА, ОБЪЕКТЫ НЕЧИСЛОВОЙ ПРИРОДЫ, ВЕРОЯТНОСТНО-СТАТИСТИЧЕСКИЕ МОДЕЛИ, ДИХОТОМИЧЕСКИЕ ДАННЫЕ, БИНАРНЫЕ ОТНОШЕНИЯ, РАНГИ, ОБЪЕКТЫ ОБЩЕЙ ПРИРОДЫ

Keywords: MATHEMATICAL STATISTICS, NONNUMERICAL OBJECTS, PROBABILISTIC-STATISTICAL MODELS, DICHOTOMOUS DATA, BINARY RELATIONS, RANKS, OBJECTS OF GENERAL NATURE

1. Введение

Статистика объектов нечисловой природы (статистика нечисловых данных, нечисловая статистика) является областью математической статистики, посвященной методам анализа нечисловых данных. Развитие этой области статистики рассмотрено в статье [1]. Многообразие нечисловых данных представлено в работе [2]. Основой применения результатов математической статистики являются вероятностно-статистические модели реальных явлений и процессов, важнейшей (а часто и единственной) составной частью которых являются модели порождения данных. Простейшим примером модели порождения данных является модель выборки как совокупности независимых одинаково распределенных случайных величин. В настоящей статье рассмотрим основные вероятностные модели порождения нечисловых данных. А именно, модели дихотомических данных, результатов парных сравнений, бинарных отношений, рангов, объектов общей природы. Обсудим различные варианты вероятностных моделей, в том числе разработанные автором и впервые представленные в научной периодике, и их практическое использование.

2. Дихотомические данные

Рассмотрим базовую вероятностную модель дихотомических данных - *бернуллиевский вектор* (в терминологии энциклопедии [3] - *люсиан*), т.е.

конечную последовательность $X = (X_1, X_2, ..., X_k)$ независимых испытаний Бернулли X_i, для которых $P(X_i = 1) = p_i$ и $P(X_i = 0) = p_i$, $i = 1,2,...,k$, причем вероятности p_i могут быть различны.

Бернуллиевские вектора часто применяются при практическом использовании статистических методов. Так, они использованы в монографии [4] для описания равномерно распределенных случайных толерантностей. Как известно, толерантность на множестве из m элементов можно задать симметричной относительно главной диагонали матрицей $\|\delta_{ij}\|$ порядка m из 0 и 1, на главной диагонали которой стоят 1. Тогда случайная толерантность описывается распределением $m(m\text{-}1)/2$ дихотомических случайных величин δ_{ij}, $1 \le i < j \le m$, а для равномерно распределенной (на множестве всех толерантностей) толерантности эти случайные величины, как можно доказать, оказываются независимыми и принимают значения 0 и 1 с равными вероятностями 1/2. Записав элементы δ_{ij} задающей такую толерантность матрицы в строку, получим бернуллиевский вектор с $k = m(m\text{-}1)/2$ и $p_i = 1/2$, $i = 1,2,...,k$.

В связи с оцениванием по статистическим данным функции принадлежности нечеткой толерантности в 1970-е годы нами была построена теория случайных толерантностей с такими независимыми δ_{ij}, что вероятности $P(\delta_{ij} = 1) = p_{ij}$ произвольны [4]. Случайные множества с независимыми элементами использовались как общий язык для описания парных сравнений и случайных толерантностей. В некоторых публикациях термин «люсиан» применялся как сокращение для выражения «случайные множества с независимыми элементами».

Был выявлен ряд областей, в которых полезен математический аппарат решения различных статистических задач, связанных с бернуллиевскими векторами. Перечислим эти области, включая ранее названные:

- анализ случайных толерантностей;

- случайные множества с независимыми элементами;

- обработка результатов независимых парных сравнений;

- статистические методы анализа точности и стабильности технологических процессов;

- анализ и синтез планов статистического приемочного контроля (по альтернативным, т.е. дихотомическим, признакам);

- обработка маркетинговых и социологических анкет (с закрытыми вопросами типа «да» - «нет»);

- обработка социально-психологических и медицинских данных, в частности, ответов на психологические тесты типа MMPI (используемых, в частности, в задачах управления персоналом),

- анализ топографических карт (применяемых для анализа и прогноза зон поражения при технологических авариях, распространении коррозии, распространении экологически вредных загрязнений, различных заболеваниях (в частности, при инфаркте миокарда [5]), в других ситуациях), и т.д.

Теорию бернуллиевских векторов можно выразить в терминах любой из этих теоретических и прикладных областей. Однако терминология одной из этих областей «режет слух» и может приводить к недоразумениям в другой из них. Поэтому целесообразно использовать термин «бернуллиевский вектор» в указанном выше значении, не связанном ни с какой из перечисленных областей приложения этой теории (в ряде публикаций в том же значении использовался термин «люсиан»).

Распределение бернуллиевского вектора X полностью описывается векторным параметром $P=(p_1, p_2,..., p_k)$, т.е. нечетким подмножеством множества $\{1, 2, ..., k\}$. Действительно, для любого детерминированного вектора $x=(x_1, x_2,..., x_k)$ из 0 и 1 имеем

$$P(X = x) = \prod_{1 \le j \le k} h(x_j, p_j),$$

где $h(x, p)=p$ при $x = 1$ и $h(x, p)=1 - p$ при $x=0$.

Теперь можно уточнить способы использования люсианов в прикладной статистике. Бернуллиевскими векторами можно моделировать:

- результаты статистического контроля (0 - годное изделие, 1 - дефектное);

- результаты маркетинговых и социологических опросов (0 - опрашиваемый выбрал первую из двух подсказок, 1 - вторую);

- распределение посторонних включений в материале (0 - нет включения в определенном объеме материала, 1 - есть);

- результаты испытаний и анализов (0 - нет нарушений требований нормативно-технической документации, 1 - есть такие нарушения);

- процессы распространения, например, пожаров (0 - нет загорания, 1 - есть; подробнее см. [4, с.215-223]);

- состояние технологического процесса (0 - процесс находится в границах допуска, 1 - вышел из них);

- ответы экспертов (опрашиваемых) о сходстве объектов (проектов, образцов), и т.д.

Математико-статистическая теория люсианов развита в статье [6].

3. Парные сравнения

Общую модель парных сравнений опишем согласно монографии Г. Дэвида [7, с.9]. Предположим, что t объектов $A_1, A_2, ..., A_t$ сравниваются попарно каждым из n экспертов. Всего возможных пар для сравнения имеется $s = t(t-1)/2$. Эксперт с номером γ делает r_γ повторных сравнений для каждой из s возможностей. Пусть $X(i, j, \gamma, \delta)$, $i,j=1,2,...,t$, $i \neq j$, $\gamma=1,2,...,n$; $\delta=1,2,...,$ r_γ, - случайная величина, принимающая значение 1 или 0 в зависимости от того, предпочитает ли эксперт с номером γ объект

A_i или объект A_j в δ-м сравнении двух объектов. Предполагается, что все сравнения проводятся независимо друг от друга, так что случайные величины $X(i, j, \gamma, \delta)$ независимы в совокупности, если не считать того, что $X(i, j, \gamma, \delta) + X(j, i, \gamma, \delta) = 1$. Положим

$$P(X(i, j, \gamma, \delta) = 1) = \pi(i, j, \gamma, \delta).$$

Ясно, что описанная модель парных сравнений представляет собой частный случай бернуллиевского вектора. В этой модели число наблюдений равно числу неизвестных параметров, поэтому для получения статистических выводов необходимо наложить априорные условия на вероятности $\pi(i, j, \gamma, \delta)$, например [7, с.9]:

$\pi(i, j, \gamma, \delta) = \pi(i, j, \gamma)$ (нет эффекта от повторений);

$\pi(i, j, \gamma, \delta) = \pi(i, j)$ (нет эффекта от повторений и от экспертов).

Теорию независимых парных сравнений целесообразно разделить на две части - непараметрическую, в которой статистические задачи ставятся непосредственно в терминах $\pi(i, j, \gamma, \delta)$, и параметрическую, в которой вероятности $\pi(i, j, \gamma, \delta)$ выражаются через меньшее число иных параметров. Ряд результатов непараметрической теории парных сравнений непосредственно вытекает из теории бернуллиевских векторов.

В параметрической теории парных сравнений наиболее популярна так называемая линейная модель [7, с.11], в которой предполагается , что каждому объекту A_i можно сопоставить некоторую «ценность» V_i так, что вероятность предпочтения $\pi(i, j)$ (т.е. предполагается дополнительно, что эффект от повторений и от экспертов отсутствует) выражается следующим образом:

$$\pi(i, j) = H(V_i - V_j), \qquad (1)$$

где $H(x)$ - функция распределения, симметричная относительно 0, т.е.

$$H(-x) = 1 - H(x) \qquad (2)$$

при всех x.

Широко применяются модели Терстоуна - Мостеллера и Брэдли - Терри, в которых $H(x)$ - соответственно функции нормального и логистического распределений. Поскольку функция $\Phi(x)$ стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1 и функция

$$\Psi(x) = e^x (1 + e^x)^{-1}$$

стандартного логистического распределения удовлетворяют (см., например, [8]) соотношению

$$\sup_{x \in R^1} | \Phi(x) - \Psi(1{,}7x) | < 0{,}01,$$

то для обоснованного выбора по статистическим данным между моделями Терстоуна - Мостеллера и Брэдли - Терри необходимо не менее тысячи наблюдений [9].

Соотношение (1) вытекает из следующей модели поведения эксперта: он измеряет «ценность» V_i и V_j объектов A_i и A_j, но с ошибками ε_i и ε_j соответственно, а затем сравнивает свои оценки ценности объектов $y_i = V_i + \varepsilon_i$ и $y_j = V_j + \varepsilon_j$. Если $y_i > y_j$, то он предпочитает A_i, в противном случае - A_j. Тогда

$$\pi(i, j) = P(\varepsilon_i - \varepsilon_j < V_i - V_j) = H(V_i - V_j). \qquad (3)$$

Обычно предполагают, что субъективные ошибки эксперта ε_i и ε_j независимы и имеют одно и то же непрерывное распределение. Тогда функция распределения $H(x)$ из соотношения (3) непрерывна и удовлетворяет функциональному уравнению (2).

Существует много разновидностей моделей парных сравнений, постоянно предлагаются новые. В качестве примера опишем разработанную нами модель парных сравнений, основанную не на процедуре упорядочения, а на определении сходства объектов. Пусть каждому объекту A_i соответствует точка a_i в r-мерном евклидовом пространстве R^r. Эксперт «измеряет» a_i и a_j с ошибками ε_i и ε_j соответственно и в случае, если евклидово расстояние между $a_i + \varepsilon_i$ и $a_j + \varepsilon_j$ меньше 1, заявляет о сходстве объектов A_i и A_j, в противном случае – об их различии. Предполагается, что ошибки ε_i и ε_j независимы и имеют одно и то же распределение, например, круговое нормальное распределение с нулевым математическим ожиданием и дисперсией координат σ^2. Целью статистической обработки является определение по результатам парных сравнений оценок параметров a_1, a_2,…,a_r, и σ^2, а также проверка согласия опытных данных с моделью.

Рассмотренные модели парных сравнений могут быть обобщены в различных направлениях. Так, можно ввести понятие «ничья» – ситуации, когда эксперт оценивает объекты одинаково. Модели с учетом «ничьих» предполагают, что эксперт может отказаться от выбора одного из объектов и заявить об их эквивалентности, т. е. число возможных ответов увеличивается с 2 до 3. В моделях множественных сравнений эксперту представляется не два объекта, а три или большее число

Модели, учитывающие «ничьи», строятся обычно с помощью используемых в психофизике «порогов чувствительности»: если $| y_i - y_j | \leq r$ (где r - порог чувствительности), то объекты A_i и A_j эксперт объявляет неразличимыми. Приведем пример разработанной нами модели с «ничьими», основанной на другом принципе. Пусть каждому объекту A_i соответствует точка a_i в r-мерном линейном пространстве. Как и прежде, эксперт «измеряет объектные точки» a_i и a_j с ошибками ε_i и ε_j соответственно, т.е. принимает решение на основе $y_i = a_i + \varepsilon_i$ и $y_j = a_j + \varepsilon_j$. Если все координаты y_i больше соответствующих координат y_j , то A_i предпочитается A_j. Соответственно, если каждая координата y_i меньше координаты y_j с тем же номером, то эксперт считает наилучшим объект A_j. Во всех остальных случаях эксперт объявляет о ничейной ситуации. Эта модель при r = 1 переходит в описанную выше линейную модель. Она связана с принципом Парето в теории группового выбора и предусматривает выбор оптимального по Парето объекта, если он существует (роль согласуемых критериев играют процедуры сравнения значений отдельных координат), и отказ от выбора, если такого объекта нет.

Можно строить модели, учитывающие порядок предъявления объектов при сравнении, зависимость результата сравнения от результатов предшествующих сравнений. Опишем одну из подобных моделей, разработанных нами.

Пусть эксперт сравнивает три объекта – *A*, *B*, *C*, причем сначала сравниваются *A* и *B*, потом – *B* и *C* и, наконец, *A* и *C*. Для определенности пусть $A > B$ будет означать, что *A* более предпочтителен, чем *B*. Пусть при предъявлении двух объектов

$$P(A > B) = \pi_{AB}, P(B > C) = \pi_{BC}, P(A > C) = \pi_{AC}.$$

Теперь пусть пара *B*, *C* предъявляется после пары *A*, *B*. Естественно предположить, что высокая оценка *B* в первом сравнении повышает вероятность предпочтения *B* и во втором, и, наоборот, отрицательное мнение о *B* в первом сравнении сохраняется и при проведении второго сравнения. Это предположение проще всего учесть в модели следующим образом:

$$P(B > C \mid B > A) = \pi_{BC} + \delta, \quad P(B > C \mid A > B) = \pi_{BC} - \delta,$$

где δ - некоторое положительное число, показывающее степень влияния первого сравнения на второе. По аналогичным причинам вероятности исхода третьего сравнения в зависимости от результатов первых двух можно описать так:

$$P(A > C \mid A > B, B > C) = \pi_{AC} + 2\delta, \quad P(A > C \mid A > B, B < C) = \pi_{AC},$$

$$P(A > C \mid A < B, B > C) = \pi_{AC}, \quad P(A > C \mid A < B, B < C) = \pi_{AC} - 2\delta.$$

Статистическая задача состоит в определении параметров π_{AB}, π_{BC}, π_{AC} и δ по результатам сравнений, проведенных n экспертами, и в проверке адекватности модели.

Ясно, что можно рассматривать и другие модели, в частности, учитывающие тягу экспертов к транзитивности ответов. Очевидно, что проблемы построения моделей парных сравнений относятся не к статистике объектов нечисловой природы, а к тем прикладным областям, для решения задач которых развиваются методы парных сравнений, например, к организации машиностроительного производства, экономике предприятия, стратегическому менеджменту, производственной психологии, изучению поведения потребителей, экспертным оценкам и т.д.

Метод парных сравнений был введен в 1860 г. Г. Т. Фехнером для решения задач психофизики. Расскажем об этом несколько подробнее. Как известно, основателем психофизики по праву считается Густав Теодор Фехнер (1801 – 1887), а год выхода в свет его фундаментальной работы «Элементы психофизики» (1860) – датой рождения новой науки. В этой работе широко применялся предложенный Г.Т. Фехнером метод парных сравнений (обсуждение событий тех лет с современных позиций дано в монографии [7, с.14-16]).

С точки зрения математической статистики приведенные выше модели не представляют большого теоретического интереса: оценки параметров находятся обычно методом максимального правдоподобия или асимптотически эквивалентным ему методом одношаговых оценок (см. [9]), а проверка согласия проводится по критерию отношения правдоподобия или асимптотически эквивалентными ему критериями типа хи-квадрат [7]. При этом вычислительные процедуры обычно достаточно сложны и плохо исследованы.

Отметим некоторые сложности при обосновании возможности использовании линейных моделей типа (1) - (3). Вероятностностатистическая теория достаточно проста, когда предполагается, что каждому отдельному сравнению двух объектов соответствуют свои собственные ошибки экспертов, причем все ошибки независимы в совокупности. Однако это предположение отнюдь не очевидно с содержательной точки зрения. В качестве примера рассмотрим три объекта *А*, *В* и *С*, которые сравнивают попарно: *А* и *В*, *В* и *С*, *А* и *С*. В соответствии со сказанным, в рассмотрение вводят 6 ошибок одного и того же эксперта: ε_A и ε_B в первом сравнении, $\varepsilon_B^{'}$ и ε_C – во втором, $\varepsilon_A^{'}$ и $\varepsilon_C^{'}$ – в третьем, причем все эти 6 случайных величин независимы в совокупности. Между тем естественно думать, что мнения эксперта об одном и том же объекте связаны между собой. Т. е. ε_A и $\varepsilon_A^{'}$ зависимы, равно как ε_B и $\varepsilon_B^{'}$, а также ε_C и $\varepsilon_C^{'}$. Более того, если принять, что точка зрения эксперта полностью определена для него самого, то следует положить $\varepsilon_A = \varepsilon_A^{'}$ и соответственно $\varepsilon_B = \varepsilon_B^{'}$ и $\varepsilon_C = \varepsilon_C^{'}$. При этом, напомним, случайные величины ε_A, ε_B и др. интерпретируется как отклонения мнений отдельных экспертов от истины. Видимо, ошибку эксперта целесообразно считать состоящей из двух слагаемых, а именно: отклонения от истины, вызванного внутренними особенностями эксперта (систематическая погрешность) и колебания мнения эксперта в связи с очередным парным сравнением (случайная погрешность). Игнорирование систематической погрешности облегчает развитие математико-статистической теории, а ее учет приводит к необходимости изучения зависимых парных сравнений.

При обработке результатов парных сравнений первый этап - проверка согласованности. Понятие согласованности уточняется различными способами, но все они имеют один и тот же смысл проверки однородности обрабатываемого материала, т.е. того, что целесообразно агрегировать мнения отдельных экспертов, объединить данные и совместно их обрабатывать. При отсутствии однородности данные разбиваются на группы (классы, кластеры, таксоны) с целью обеспечения однородности внутри отдельных групп. Естественно, согласованность целесообразно проверять, вводя возможно меньше гипотез о структуре данных. Следовательно, целесообразно пользоваться для этого непараметрической теорией парных сравнений, основанной на теории бернуллиевских векторов.

Хорошо известно, что модели парных сравнений с успехом применяются в экспертных и экспериментальных процедурах упорядочивания и выбора. В частности, для анализа голосований, турниров, выбора наилучшего объекта (проекта, образца, кандидатуры); в планировании и анализе сравнительных экспериментов и испытаний; в органолептической экспертизе (в частности, дегустации); при изучении поведения потребителей; визуальной колоритмии (принятии решений на основе цвета), определении индивидуальных рейтингов и вообще изучении предпочтений при выборе и т. д. (подробнее см. [4, 7]).

4. Бинарные отношения

Теорию ранговой корреляции [10] можно рассматривать как теорию статистического анализа случайных ранжировок, равномерно распределенных на множестве всех ранжировок. Так, при обработке данных классического психофизического эксперимента по упорядочению кубиков соответственно их весу, подробно описанного в работе [11], оказалась адекватной следующая т.н. *T*-модель ранжирования.

Пусть имеется t объектов $A_1, A_2, ..., A_t$, причем каждому объекту A_i соответствует число a_i, описывающее его положение на шкале изучаемого признака. Испытуемый упорядочивает объекты так, как если бы оценивал соответствующие им значения с ошибками, т.е. находил $y_i = a_i + \varepsilon_i$, i = 1, 2, …, n, где ε_i - ошибка при рассмотрении i-го объекта, а затем располагал бы объекты в том порядке, в каком располагаются $y_1, y_2,..., y_t$. В этом случае вероятность появления упорядочения $A_{i1}, A_{i2},...A_{it}$ есть $P(y_{i1} < y_{i2} < ... < y_{it})$, а ранги $R_1, R_2,..., R_t$ объектов являются рангами случайных величин $y_1, y_2,..., y_t$, полученными при их упорядочении в порядке возрастания. Кроме того, для простоты расчетов в модели предполагается, что ошибки испытуемого $\varepsilon_1, \varepsilon_2,..., \varepsilon_t$ независимы и имеют нормальное распределение с математическим ожиданием 0 и дисперсией σ^2.

Как уже отмечалось, бинарное отношение на множестве из t элементов полностью описывается матрицей из 0 и 1 порядка $t \times t$. Поэтому задать распределение случайного бинарного отношения - это то же самое, что задать распределение вероятностей на множестве всех матриц описанного вида, состоящем из $2^{(t^2)}$ элементов. Пространства ранжировок, разбиений, толерантностей зачастую удобно считать подпространствами пространства всех бинарных отношений, тогда распределения вероятностей на них - частные случаи описанного выше распределения, выделенные тем, что вероятности принадлежности случайного бинарного отношения соответствующим подпространствам равны 1. Распределение произвольного бинарного отношения описывается $2^{(t^2)} - 1$ параметрами, распределение случайной ранжировки (без связей) – (t! – 1) параметрами, а описанная выше T-модель ранжирования задается (t + 1) параметрами. При t = 4 эти числа равны соответственно 65535, 23 и 5. Первое из этих чисел показывает практическую невозможность использования произвольных бинарных отношений в предназначенных для практического использования вероятностно-статистических моделях, поскольку по имеющимся данным невозможно оценить столь большое число параметров. Приходится ограничиваться теми или иными семействами бинарных отношений - ранжировками, разбиениями, толерантностями и др. Модель произвольной случайной ранжировки при t = 5 описывается

119 параметрами, при t = 6 - уже 719 параметрами, при t = 7 число параметром достигает 5049, что уже явно за пределами возможности оценивания. В то же время T-модель ранжирования при t = 7 описывается всего 8-ю параметрами, а потому может быть кандидатом для практического использования.

Что естественно предположить относительно распределения случайного элемента со значениями в том или ином пространстве бинарных отношений? Зачастую целесообразно считать, что распределение имеет некий центр, попадание в который наиболее вероятно, а по мере удаления от центра вероятности убывают. Это соответствует естественной модели измерения с ошибкой. В классическом одномерном случае результат подобного измерения обычно описывается унимодальной симметричной плотностью, монотонно возрастающей слева от модального значения, в котором плотность максимальна, и монотонно убывающей справа от него. Чтобы ввести понятие монотонного распределения в пространстве бинарных отношений, будем исходить из метрики в этом пространстве. Воспользовавшись тем, что бинарные отношения C и D однозначно описываются матрицами $\| c_{ij} \|$ и $\| d_{ij} \|$ порядка $t \times t$, рассмотрим расстояние (в несколько другой терминологии - метрику) в пространстве бинарных отношений

$$d(C,D) = \sum_{1 \le i,\, j \le t} | c_{ij} - d_{ij} |. \qquad (4)$$

Метрика (4) в различных пространствах бинарных отношений - ранжировок, разбиений, толерантностей - может быть введена с помощью соответствующих систем аксиом [12]. В настоящее время метрику (4) обычно называют расстоянием Кемени в честь американского исследователя Джона Кемени, впервые получившего эту метрику исходя из предложенной им системы аксиом для расстояния между упорядочениями (ранжировками).

В статистике нечисловых данных используются и иные метрики, отличающиеся от расстояния Кемени. Более того, для использования понятия монотонного распределения, о котором сейчас идет речь, нет необходимости требовать выполнения неравенства треугольника, а достаточно, чтобы $d(C,D)$ можно было рассматривать как показатель различия. Под показателем различия понимаем такую функцию $d(C,D)$ двух бинарных отношений C и D, что $d(C,D) = 0$ при $C = D$ и увеличение $d(C,D)$ интерпретируется как возрастание различия между C и D.

Определение 1. Распределение бинарного отношения X называется монотонным с центром в C_0 относительно расстояния (показателя различия) d, если из $d(C,C_0) < d(D,C_0)$ следует, что $P(X=C) > P(X=D)$.

Это определение впервые введено в монографии [4, с.196]. Оно может использоваться в любых пространствах бинарных отношений и, более того, в любых пространствах из конечного числа элементов, лишь бы в них была введена функция $d(C,D)$ - показатель различия элементов C и D этого пространства. Монотонное распределение унимодально, мода находится в C_0.

Определение 2. Распределение бинарного отношения X называется симметричным относительно расстояния d с центром в C_0, если существует такая функция $f: R_+^1 \to [0,1]$, что

$$P(X = C) = f(d(C, C_0)). \qquad (5)$$

Если распределение X монотонно и таково, что из $d(C,C_0) = d(D,C_0)$ следует $P(X = C) = P(X = D)$, то оно симметрично. Если функция f в формуле (5) монотонно строго убывает, то соответствующее распределение монотонно в смысле определения 1.

Поскольку толерантность на множестве из t элементов задается $0{,}5t(t - 1)$ элементами матрицы из 0 и 1 порядка $t \times t$, лежащими выше главной диагонали, то всего толерантностей имеется $2^{0,5t(t-1)}$, а потому распределение на множестве толерантностей задается в общем случае

$2^{0,5t(t-1)}$-1 параметрами. Естественно выделить семейство распределений, соответствующее независимым элементам матрицы. Оно задается бернуллиевским вектором (люсианом) с $0,5t(t-1)$ параметрами (выше бернуллиевские вектора рассмотрены подробнее). Математическая техника, необходимая для изучения толерантностей с независимыми элементами, существенно проще, чем в случае ранжировок и разбиений. Здесь легко отказаться от условия равномерности распределения. Этому условию соответствует $p_{ij} \equiv 1/2$, в то время как статистические методы анализа люсианов, развитые в статистике нечисловых данных (см., например, работы [4, 6, 13, 14]) не налагают никаких существенных ограничений на p_{ij} .

Как уже отмечалось, при обработке мнений экспертов сначала проверяют согласованность. В частности, если мнения экспертов описываются монотонными распределениями, то для согласованности необходимо совпадение центров этих распределений. К сожалению, рассмотренные выше классические методы проверки согласованности для ранжировок, основанные на коэффициентах ранговой корреляции и конкордации, позволяют лишь отвергнуть гипотезу о равнораспределенности. Но не установить, можно ли считать, что центры соответствующих экспертам распределений совпадают или же, например, существует две группы экспертов, каждая со своим центром. Теория случайных толерантностей лишена этого недостатка. Отсюда вытекают следующие практические рекомендации.

Пусть цель обработки экспертных данных состоит в получении ранжировки, отражающей групповое мнение. Однако согласно рекомендуемой процедуре экспертного опроса пусть эксперты не упорядочивают объекты, а проводят парные сравнения, сравнивая каждый из рассматриваемых объектов со всеми остальными, причем ровно один раз. Тогда ответ эксперта - толерантность, но, вообще говоря, не ранжировка, поскольку в ответах эксперта может нарушаться транзитивность.

Возможны два пути обработки данных. Первый - превратить ответ эксперта в ранжировку (тем или иным способом «спроектировав» его на пространство ранжировок), а затем проверять согласованность ранжировок с помощью известных критериев. При этом от толерантности перейти к ранжировке можно, например, так. Будем выбирать ближайшую (в смысле применяемого расстояния) матрицу к матрице ответов эксперта из всех, соответствующих ранжировкам без связей.

Второй путь - проверить согласованность случайных толерантностей, а групповое мнение искать с помощью медианы Кемени непосредственно по исходным данным, т.е. по толерантностям. Групповое мнение при этом может быть найдено в пространстве ранжировок. Второй путь мы считаем более предпочтительным, поскольку при этом обеспечивается более адекватная проверка согласованности и исключается процедура укладывания мнения эксперта в «прокрустово ложе» ранжировки (эта процедура может приводить как к потере информации, так и к принципиально неверным выводам, вызванным искажениями мнений экспертов).

Для бинарного отношения с монотонным симметричным распределением теоретическое среднее (математическое ожидание) совпадает с центром распределения, а потому в силу закона больших чисел медиана Кемени при росте объема выборки (числа экспертов) стремится к центру распределения [15].

Области применения статистики бинарных отношений многообразны: ранговая корреляция - оценка величины связи между переменными, измеренными в порядковой шкале; анализ экспертных или экспериментальных упорядочений; анализ разбиений технико-экономических показателей на группы сходных между собой; обработка данных о сходстве (взаимозаменяемости); статистический анализ классификаций; математические вопросы теории менеджмента и др.

5. Случайные множества

Будем рассматривать случайные подмножества некоторого множества *Q*. Если *Q* состоит из конечного числа элементов, то считаем, что случайное подмножество *S* - это случайный элемент со значениями в 2^Q - множестве всех подмножеств множества *Q*, состоящем из $2^{card(Q)}$ элементов. Чтобы удовлетворить требованиям математической строгости, примем, что все подмножества *Q* измеримы (другими словами, σ-алгебра измеримых множеств совпадает с совокупностью всех подмножеств рассматриваемого конечного множества). Тогда распределение случайного подмножества $S = S(\omega)$ множества *Q* – это

$$P_S(A) = P(S = A) = \mathrm{P}(\{\omega : S(\omega) = A\}), A \subseteq Q. \qquad (6)$$

В формуле (6) предполагается, что $S : \Omega \to 2^Q$, где (Ω, F, P) - вероятностное пространство, на котором определен случайный элемент $S(\omega)$. (Здесь Ω - пространство элементарных событий, *F* - σ-алгебра случайных событий, P - вероятностная мера на *F*.) Через распределение $P_S(A)$ выражаются вероятности различных событий, связанных с *S*. Так, чтобы найти вероятность накрытия фиксированного элемента *q* случайным множеством *S*, достаточно вычислить

$$P(q \in S) = \mathrm{P}(\{\omega : q \in S(\omega)\}) = \sum_{A: q \in A, A \subseteq 2^Q} P(S = A),$$

где суммирование идет по всем подмножествам *A* множества *Q*, содержащим *q*. Пусть $Q=\{q_1, q_2, \dots, q_k\}$. Рассмотрим случайные величины, определяемые по случайному множеству *S* следующим образом

$$\chi_i(\omega) = \begin{cases} 1, q_i \in S(\omega), \\ 0, q_i \notin S(\omega). \end{cases}$$

Определение 3. Случайное множество S называется случайным множеством с независимыми элементами, если случайные величины $\chi_i(\omega), i = 1,2,...,k,$ независимы (в совокупности).

Последовательность случайных величин $\chi_1, \chi_2, ..., \chi_k$ --бернуллиевский вектор с $X_i = \chi_i$ и $p_i = P(q_i \in S(\omega)), i = 1,2,...,k.$ Из сказанного выше следует, что распределение случайного множества с независимыми элементами задается формулой

$$P(S = A) = \prod_{q_i \in A} p_i \prod_{q_i \in Q \setminus A} (1 - p_i),$$

т.е. такие распределения образуют параметрическое семейство размерности $k = card(Q)$, входящее в семейство всех распределений случайных подмножеств множества Q размерности $(2^{card(Q)} - 1)$.

При исследовании случайных подмножеств произвольного множества Q будем рассматривать их как случайные величины со значениями в некотором пространстве подмножеств множества Q, например, в пространстве замкнутых подмножеств 2^Q множества Q.

Представляющими интерес лишь для математиков способами введения измеримой структуры в 2^Q интересоваться не будем. Отсутствие специального интереса к проблеме измеримости связано с тем, что при вероятностно-статистическом моделировании и обработке на ЭВМ все случайные подмножества необходимо рассматривать как конечные (т.е. подмножества конечного множества).

Случайные множества находят разнообразные применения в многообразных проблемах эконометрики и математической экономики. В том числе в задачах управлении запасами и ресурсами (см. об этом главу 5 в монографии [4), в задачах менеджмента и, в частности, маркетинга, в экспертных оценках, например, при анализе мнений голосующих или опрашиваемых, каждый из которых отмечает несколько пунктов из списка и т.д. Кроме того, случайные множества применяются в гранулометрии, при изучении пористых сред и объектов сложной природы в таких областях, как металлография, петрография, биология, в частности, математическая морфология. Они используются при изучении структуры веществ и материалов, в исследовании процессов распространения, в том числе просачивания, распространения пожаров, экологических загрязнений, при районировании, в изучении областей поражения, например, поражения металла коррозией и сердечной мышцы при инфаркте миокарда, и т.д., и т.п. Можно вспомнить о компьютерной томографии, о наглядном представлении сложной информации на экране компьютера, об изучении распространения рекламной информации, о картах Кохонена (популярный метод представления информации при применении нейросетей) и т.д.

6. Ранговые методы

Ранее в настоящей статье установлено, что любой адекватный алгоритм в порядковой шкале является функцией от некоторой матрицы C. Пусть никакие два из результатов наблюдений $x_1, x_2,\dots,x_n$ не совпадают, а $r_1, r_2,\dots,r_n$ - их ранги. Тогда элементы матрицы C и ранги результатов наблюдений связаны взаимно однозначным соответствием:

$$r_i = 1 + \sum_{1 \le j \le n} (1 - c_{ij}),$$

а c_{ij} через ранги выражаются так: $c_{ij} = 1$, если $r_i < r_j$, и $c_{ij} = 0$ в противном случае.

Сказанное означает, что при обработке данных, измеренных в порядковой шкале, могут применяться только ранговые статистические методы (т.е. все используемые статистики зависят только от рангов результатов измерений, а не от их численных величин). Отметим, что часто используемое в непараметрической статистике преобразование Смирнова $Y = F(X)$ (здесь $F(x)$ - непрерывная функция распределения случайной величины X, причем F предполагается произвольной) фактически означает переход к порядковой шкале, поскольку статистические выводы при этом инвариантны относительно допустимых преобразований в порядковой шкале.

Разумеется, ранговые статистические методы могут применяться не только при обработке данных, измеренных в порядковой шкале. Так, для проверки независимости двух количественных признаков в случае, когда нет уверенности в нормальности соответствующего двумерного распределения, целесообразно пользоваться коэффициентами ранговой корреляции Кендалла или Спирмена [10] (отметим, что используют два варианта написания фамилии Kendall на русском языке – Кендалл (издательства «Наука» и «Мир») и Кендэл (издательство «Статистика»), что иногда может приводить к недоразумениям).

В настоящее время с помощью непараметрических и прежде всего ранговых методов можно решать тот же набор задач прикладной статистики, что и с помощью параметрических методов, в частности, основанных на предположении нормальности. Однако параметрические методы вошли в массовое сознание исследователей и инженеров и мешают широкому внедрению более обоснованной и прогрессивной ранговой статистики. Так, при проверке однородности двух выборок вместо критерия Стьюдента целесообразно использовать ранговые методы [9], но пока это делается редко.

7. Объекты общей природы

Вероятностная модель объекта нечисловой природы в общем случае - случайный элемент со значениями в пространстве произвольного вида, а модель выборки таких объектов - совокупность независимых одинаково распределенных случайных элементов. Именно такая модель была использована для обработки наблюдений, каждое из которых - нечеткое множество [16, 17].

Из-за имеющего разнобоя в терминологии приведем математические определения из справочника по теории вероятностей академика РАН Ю.В Прохорова и проф. Ю.А. Розанова [18] (см. также [3]).

Пусть (X, B) - некоторое измеримое пространство; (F, B)-измеримая функция $\xi = \xi(\omega)$ на пространстве элементарных событий (Ω, F, P) (где P - вероятностная мера на σ-алгебре F - измеримых подмножеств Ω, называемых событиями) со значениями в (X, B) называется случайной величиной (чаще этот математический объект называют случайным элементом, оставляя термин «случайная величина» за частным случаем, когда X - числовая прямая - А.О.) в фазовом пространстве (X, B). Распределением вероятностей этой случайной величины ξ называется функция $P_\xi = P_\xi(B)$ на σ-алгебре B фазового пространства, определенная как

$$P_\xi = P\{\xi \in B\} \quad (B \in B) \qquad (7)$$

(распределение вероятностей P_ξ представляет собой вероятностную меру в фазовом пространстве (X, B)) [18, с. 132].

Пусть $\xi_1, \xi_2, ..., \xi_n$ - случайные величины на пространстве случайных событий (Ω, F, P) в соответствующих фазовых пространствах (X_k, B_k). Совместным распределением вероятностей этих величин называется функция $P_{\xi_1,\xi_2,...,\xi_n} = P_{\xi_1,\xi_2,...,\xi_n}(B_1, B_2, ..., B_n)$, определенная на множествах $B_1 \in B_1$, $B_2 \in B_2$, $\ldots, B_n \in B_n$ как

$$P_{\xi_1,\xi_2,...,\xi_n}(B_1, B_2, ..., B_n) = P_{\xi_1,\xi_2,...,\xi_n}(\xi_1 \in B_1, \xi_2 \in B_2, ..., \xi_n \in B_n). \qquad (8)$$

Распределение вероятностей $P_{\xi_1,\xi_2,...,\xi_n}$ как функция на полукольце множеств вида $B_1 \times B_2 \times ... \times B_n$, $B_1 \in B_1, B_2 \in B_2, ..., B_n \in B_n$, в произведении пространств $X_1, X_2, ..., X_n$ представляет собой функцию распределения.

Случайные величины $\xi_1,\xi_2,...,\xi_n$ называются независимыми, если при любых B_1, B_2,…,B_n (см. [18, с.133])

$$P_{\xi_1,\xi_2,...,\xi_n}(B_1,B_2,...,B_n)=P_{\xi_1}(B_1)P_{\xi_2}(B_2)...P_{\xi_n}(B_n)\,. \qquad (9)$$

Предположим, что совместное распределение вероятностей $P_{\xi,\eta}(A,B)$ случайных величин ξ и η абсолютно непрерывно относительно некоторой меры Q на произведении пространств $\mathrm{X}\times\mathrm{Y}$, являющейся произведением мер Q_{X} и Q_{Y}, т.е.:

$$P_{\xi,\eta}(A,B)=\int\limits_{A\times B}p(x,y)Q(dx,dy) \qquad (10)$$

для любых $A\in \mathrm{A}$ и $B\in \mathrm{B}$, где *p(x,y)* - соответствующая плотность распределения вероятностей [18, с.145].

В формуле (10) предполагается, что $\xi=\xi(\omega)$ и $\eta=\eta(\omega)$ - случайные величины на одном и том же пространстве элементарных событий Ω со значениями в фазовых пространствах (X,A) и (Y,B). Существование плотности *p(x,y)* вытекает из абсолютной непрерывности $P_{\xi,\eta}(A,B)$ относительно *Q* в соответствии с теоремой Радона - Никодима.

Условное распределение вероятностей $P_{\xi}(A\,|\,\eta), A\in \mathrm{A},$ может быть выбрано одинаковым для всех $\omega\in\Omega,$ при которых случайная величина $\eta=\eta(\omega)$ сохраняет одно и то же значение: $\eta(\omega)=y.$ При почти каждом $y\in \mathrm{Y}$ (относительно распределения P_η в фазовом пространстве (Y,B)) условное распределение вероятностей $P_{\xi}(A\,|\,y)=P_{\omega,\xi}(A),$ где $\omega\in\{\eta=y\}$ и $A\in \mathrm{A},$ будет абсолютно непрерывно относительно меры Q_{X}:

$$Q_{\mathrm{X}}(A)=\int\limits_{A\times \mathrm{X}}Q(dx,dy).$$

Причем соответствующая плотность условного распределения вероятностей будет иметь вид (см. [18, с.145-146]):

$$p_{\xi}(x\,|\,y)=\frac{P_{\xi}(dx\,|\,y)}{Q_{\mathrm{X}}(dx)}=\frac{p(x,y)}{\int\limits_{\mathrm{X}}p(x,y)Q_{\mathrm{X}}(dx)}. \qquad (11)$$

При построении вероятностных моделей реальных явлений важны вероятностные пространства из конечного числа элементарных событий. Для них перечисленные выше общие понятия становятся более прозрачными, в частности, снимаются вопросы измеримости (все подмножества конечного множества обычно считаются измеримыми). Вместо плотностей и условных плотностей рассматриваются вероятности и условные вероятности. Отметим, что вероятности можно рассматривать как плотности относительно меры, приписывающей каждому элементу пространства элементарных событий вес 1, т.е. считающей меры

$$Q(A) = Card(A)$$

(мера каждого множества равна числу его элементов). В целом ясно, что определения основных понятий теории вероятностей в общей ситуации практически не отличаются от таковых в элементарных курсах, во всяком случае с идейной точки зрения.

За последние тридцать пять лет в математической статистике сформировалась новая область - нечисловая статистика, или статистика нечисловых данных, она же - статистика объектов нечисловой природы [1]. К настоящему времени она развита не менее, чем ранее выделенные статистика случайных величин, многомерный статистический анализ, статистика временных рядов и случайных процессов. Краткая сводка основных постановок и результатов математической статистики в пространствах нечисловой природы дана в книге [19]. Ряд вероятностных моделей конкретных видов объектов нечисловой природы рассмотрен в статье [20].

Теория, построенная для результатов наблюдений, лежащих в пространствах общей природы, является центральным стержнем в нечисловой статистике. В ее рамках удалось разработать и изучить методы оценивания параметров и характеристик, проверки гипотез (в частности, с помощью статистик интегрального типа), параметрической и непараметрической регрессии (восстановления зависимостей), непараметрического оценивания плотности, дискриминантного и кластерного анализов и т.д.

Вероятностно-статистические методы, развитые для результатов наблюдений, принадлежащих пространствам произвольного вида, позволяют единообразно проводить анализ данных из любого конкретного пространства. Так, в монографии [4] они применены к конечным случайным множествам, в работах [16, 17] - к нечетким множествам. С их помощью установлено поведение обобщенного мнения экспертной комиссии (медианы Кемени) при увеличении числа экспертов, когда ответы экспертов лежат в том или ином пространстве бинарных отношений. Математические методы классификации могут быть основаны на непараметрических оценках плотности распределения вероятностей в пространстве общей природы [21]. Такие методы были применены для медицинской диагностики в пространстве разнотипных данных, когда часть координат вектора измерена по количественным шкалам, а часть - по качественным, и т.д.

Литература

1. Орлов А.И. О развитии статистики объектов нечисловой природы / А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №09(093). С. 273 – 309. – IDA [article ID]: 0931309019. – Режим доступа: http://ej.kubagro.ru/2013/09/pdf/19.pdf
2. Орлов А.И. Многообразие объектов нечисловой природы / А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №08(102). С. 32 – 63. – IDA [article ID]: 1021408002. – Режим доступа: http://ej.kubagro.ru/2014/08/pdf/02.pdf, 2 у.п.л., импакт-фактор РИНЦ=0,346.
3. Вероятность и математическая статистика: Энциклопедия / Гл. ред. Ю.В. Прохоров. – М.: Большая Российская энциклопедия, 1999. – 910 с.
4. Орлов А.И. Устойчивость в социально-экономических моделях. – М.: Наука, 1979. – 296 с.

5. Попов В.Г., Аксенова Г.А., Орлов А.И., Розова Н.К., Кузьмина Е.С. Кинетокардиография в определении зон асинергии у больных инфарктом миокарда // Клиническая медицина. 1982. Т. LX. № 3. С. 25 – 30.
6. Орлов А.И. Теория люсианов / А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №07(101). С. 275 – 304. – IDA [article ID]: 1011407015. – Режим доступа: http://ej.kubagro.ru/2014/07/pdf/15.pdf
7. Дэвид Г. Метод парных сравнений. – М.: Статистика, 1978. – 144 с.
8. Орлов А.И. Логистическое распределение // Математическая энциклопедия. Т.3. – М.: Советская энциклопедия, 1982. – С. 414 – 414.
9. Орлов А.И. Прикладная статистика. – М.: Экзамен, 2006. – 671 с.
10. Кендэл М. Ранговые корреляции. – М.: Статистика, 1975. – 216 с.
11. Тюрин Ю.Н., Василевич А.П., Андрукович П.Ф. Статистические модели ранжирования // Статистические методы анализа экспертных оценок. Ученые записки по статистике, т.29. – М.: Наука, 1977. – С. 30 – 58.
12. Орлов А.И. Расстояния в пространствах статистических данных / А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №07(101). С. 227 – 252. – IDA [article ID]: 1011407013. – Режим доступа: http://ej.kubagro.ru/2014/07/pdf/13.pdf
13. Орлов А.И. Случайные множества с независимыми элементами (люсианы) и их применения // Алгоритмическое и программное обеспечение прикладного статистического анализа. Ученые записки по статистике, т.36. – М.: Наука, 1980. – С. 287 – 308.
14. Орлов А.И. Парные сравнения в асимптотике Колмогорова // Экспертные оценки в задачах управления. – М.: Изд-во Института проблем управления АН СССР, 1982. – С. 58 – 66.
15. Орлов А.И. Средние величины и законы больших чисел в пространствах произвольной природы / А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №05(089). С. 556 – 586. – IDA [article ID]: 0891305038. – Режим доступа: http://ej.kubagro.ru/2013/05/pdf/38.pdf
16. Орлов А.И. Задачи оптимизации и нечеткие переменные. - М.: Знание, 1980. – 64 с.
17. Орлов А.И., Луценко Е.В. Системная нечеткая интервальная математика. Монография (научное издание). – Краснодар, КубГАУ. 2014. – 600 с.
18. Прохоров Ю.В., Розанов Ю.А. Теория вероятностей. (Основные понятия. Предельные теоремы. Случайные процессы.) – М.: Наука, 1973. – 496 с.
19. 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/
20. Орлов А.И. Вероятностные модели конкретных видов объектов нечисловой природы. – Журнал «Заводская лаборатория». 1995. Т.61. No.5. С.43-51.
21. Орлов А.И. Математические методы теории классификации / А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №01(095). С. 423 – 459. – IDA [article ID]: 0951401023. – Режим доступа: http://ej.kubagro.ru/2014/01/pdf/23.pdf

References

1. Orlov A.I. O razvitii statistiki ob#ektov nechislovoj prirody / A.I. Orlov // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2013. – №09(093). S. 273 – 309. – IDA [article ID]: 0931309019. – Rezhim dostupa: http://ej.kubagro.ru/2013/09/pdf/19.pdf
2. Orlov A.I. Mnogoobrazie ob#ektov nechislovoj prirody / A.I. Orlov // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2014. – №08(102). S. 32 – 63. – IDA [article ID]: 1021408002. – Rezhim dostupa: http://ej.kubagro.ru/2014/08/pdf/02.pdf, 2 u.p.l., impakt-faktor RINC=0,346.
3. Verojatnost' i matematicheskaja statistika: Jenciklopedija / Gl. red. Ju.V. Prohorov. – M.: Bol'shaja Rossijskaja jenciklopedija, 1999. – 910 s.
4. Orlov A.I. Ustojchivost' v social'no-jekonomicheskih modeljah. – M.: Nauka, 1979. – 296 s.
5. Popov V.G., Aksenova G.A., Orlov A.I., Rozova N.K., Kuz'mina E.S. Kinetokardiografija v opredelenii zon asinergii u bol'nyh infarktom miokarda // Klinicheskaja medicina. 1982. T. LX. № 3. S. 25 – 30.
6. Orlov A.I. Teorija ljusianov / A.I. Orlov // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2014. – №07(101). S. 275 – 304. – IDA [article ID]: 1011407015. – Rezhim dostupa: http://ej.kubagro.ru/2014/07/pdf/15.pdf
7. Djevid G. Metod parnyh sravnenij. – M.: Statistika, 1978. – 144 s.
8. Orlov A.I. Logisticheskoe raspredelenie // Matematicheskaja jenciklopedija. T.3. – M.: Sovetskaja jenciklopedija, 1982. – S. 414 – 414.
9. Orlov A.I. Prikladnaja statistika. – M.: Jekzamen, 2006. – 671 s.
10. Kendjel M. Rangovye korreljacii. – M.: Statistika, 1975. – 216 s.
11. Tjurin Ju.N., Vasilevich A.P., Andrukovich P.F. Statisticheskie modeli ranzhirovanija // Statisticheskie metody analiza jekspertnyh ocenok. Uchenye zapiski po statistike, t.29. – M.: Nauka, 1977. – S. 30 – 58.
12. Orlov A.I. Rasstojanija v prostranstvah statisticheskih dannyh / A.I. Orlov // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2014. – №07(101). S. 227 – 252. – IDA [article ID]: 1011407013. – Rezhim dostupa: http://ej.kubagro.ru/2014/07/pdf/13.pdf
13. Orlov A.I. Sluchajnye mnozhestva s nezavisimymi jelementami (ljusiany) i ih primenenija // Algoritmicheskoe i programmnoe obespechenie prikladnogo statisticheskogo analiza. Uchenye zapiski po statistike, t.36. – M.: Nauka, 1980. – S. 287 – 308.
14. Orlov A.I. Parnye sravnenija v asimptotike Kolmogorova // Jekspertnye ocenki v zadachah upravlenija. – M.: Izd-vo Instituta problem upravlenija AN SSSR, 1982. – S. 58 – 66.
15. Orlov A.I. Srednie velichiny i zakony bol'shih chisel v prostranstvah proizvol'noj prirody / A.I. Orlov // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2013. – №05(089). S. 556 – 586. – IDA [article ID]: 0891305038. – Rezhim dostupa: http://ej.kubagro.ru/2013/05/pdf/38.pdf
16. Orlov A.I. Zadachi optimizacii i nechetkie peremennye. - M.: Znanie, 1980. – 64 s.

17. Orlov A.I., Lucenko E.V. Sistemnaja nechetkaja interval'naja matematika. Monografija (nauchnoe izdanie). – Krasnodar, KubGAU. 2014. – 600 s.

18. Prohorov Ju.V., Rozanov Ju.A. Teorija verojatnostej. (Osnovnye ponjatija. Predel'nye teoremy. Sluchajnye processy.) – M.: Nauka, 1973. – 496 s.

19. 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/

20. Orlov A.I. Verojatnostnye modeli konkretnyh vidov ob#ektov nechislovoj prirody. – Zhurnal «Zavodskaja laboratorija». 1995. T.61. No.5. S.43-51.

21. Orlov A.I. Matematicheskie metody teorii klassifikacii / A.I. Orlov // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2014. – №01(095). S. 423 – 459. – IDA [article ID]: 0951401023. – Rezhim dostupa: http://ej.kubagro.ru/2014/01/pdf/23.pdf