

УДК 519.2

UDC 519.2

**ТЕОРИЯ ЛЮСИАНОВ****THEORY OF LUSIANS**

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

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

*Bauman Moscow State Technical University,  
Moscow, Russia*

Люсиан – модель дихотомических данных в виде конечной последовательности независимых испытаний Бернулли с, вообще говоря, различными вероятностями успеха. Рассмотрены задачи проверки статистических гипотез, классификации, усреднения люсианов. В соответствии с потребностями практики постановки задач рассмотрены, прежде всего, в асимптотике растущей размерности, в которой число неизвестных параметров растет пропорционально объему данных. Оказались полезными метод проверки гипотез по совокупности малых выборок и теория несмещенных статистических оценок

Lusian - dichotomous data model as a finite sequence of independent Bernoulli trials with, generally speaking, different probabilities of success. The tasks of statistical hypothesis testing, classification, averaging lusians are discussed. In accordance with the requirements of practice the problems are considered primarily in the asymptotic behavior of increasing dimension, in which the number of unknown parameters increases in proportion to the size of data. The method of testing hypotheses on set of small samples and the theory of unbiased statistical estimates were proved to be useful

Ключевые слова: СТАТИСТИЧЕСКИЕ МЕТОДЫ, МАТЕМАТИЧЕСКАЯ СТАТИСТИКА, ПРИКЛАДНАЯ СТАТИСТИКА, НЕЧИСЛОВАЯ СТАТИСТИКА, ЛЮСИАНЫ, ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ, АСИМПТОТИКА РАСТУЩЕЙ РАЗМЕРНОСТИ, НЕСМЕЩЕННЫЕ СТАТИСТИЧЕСКИЕ ОЦЕНКИ, ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ, АСИМПТОТИЧЕСКАЯ НОРМАЛЬНОСТЬ

Keywords: STATISTICAL METHODS, MATHEMATICAL STATISTICS, APPLIED STATISTICS, NON-NUMERIC STATISTICS, LUSIANS, STATISTICAL HYPOTHESIS TESTING, ASYMPTOTIC BEHAVIOR OF INCREASING DIMENSION, UNBIASED STATISTICAL ESTIMATES, LIMIT THEOREMS, ASYMPTOTIC NORMALITY

## 1. Введение

В статье впервые в научной периодике систематически развивается математико-статистическая теория анализа дихотомических данных. В соответствии с потребностями практики постановки задач рассмотрены, прежде всего, в асимптотике растущей размерности, в которой число неизвестных параметров растет пропорционально объему данных. Оказалось, что для решения рассматриваемых задач полезна теория несмещенных статистических оценок.

## 2. Дихотомические (бинарные) данные

Это данные, которые могут принимать одно из двух значений (0 или 1), т.е. результаты измерений значений альтернативного признака [1]. Как

известно [2], измерения в шкале наименований и порядковой шкале приводят к бинарным отношениям, а те могут быть выражены как результаты измерений по нескольким альтернативным признакам, соответствующим элементам матриц из 0 и 1, описывающих бинарные отношения. Дихотомические данные возникают в прикладных исследованиях и многими иными путями.

В настоящее время в большинстве технических регламентов, стандартов, технических условий, договоров на поставку конкретной продукции предусмотрен контроль по альтернативному признаку. Это означает, что единица продукции относится к одной из двух категорий – «годных» или «дефектных», т.е. соответствующих или не соответствующих требованиям соответствующего нормативно-технического документа. Отечественными специалистами проведены обширные теоретические исследования проблем статистического приемочного контроля по альтернативному признаку. основополагающими в этой области являются работы академика А.Н. Колмогорова. Подход отечественной вероятностно-статистической школы к проблемам контроля качества продукции отражен в фундаментальных монографиях Ю.К. Беляева [3] и Я.П. Лумельского [4] (см. также главу 10 учебника [5]).

Дихотомические данные – давний объект математической статистики. Особенно большое применение они имеют в экономических и социологических исследованиях, в которых большинство переменных, интересующих специалистов, измеряется по качественным шкалам. При этом дихотомические данные зачастую являются более адекватными, чем результаты измерений по методикам, использующим большее число градаций. В частности, психологические тесты типа ММРІ (расшифровывается как «Миннесотское многофакторное личностное исследование») используют только дихотомические данные. На них

опираются и популярные в технико-экономическом анализе (и вообще в экспертных оценках) методы парных сравнений [6, 7].

Элементарным актом в методе парных сравнений является предъявление эксперту для сравнения двух объектов (сравнение может проводиться также прибором). В одних постановках эксперт должен выбрать из двух объектов лучший по качеству, в других - ответить, похожи объекты или нет. В обоих случаях ответ эксперта можно выразить одной из двух цифр (меток) – 0 или 1. В первой постановке: 0, если лучшим объявлен первый объект; 1 – если второй. Во второй постановке: 0, если объекты похожи, схожи, близки; 1 – в противном случае.

Подводя итоги, можно сказать, что рассмотренные выше виды данные могут быть представлены в виде векторов из 0 и 1 (при обосновании этого утверждения используется тот очевидный факт, что матрицы могут быть записаны в виде векторов). Более того, поскольку все мыслимые результаты наблюдений имеют лишь несколько значащих цифр, то, используя двоичную систему счисления, любые виды анализируемых статистическими методами данных можно записать в виде векторов конечной длины (размерности) из 0 и 1. Представляется, однако, что эта возможность в большинстве случаев имеет лишь академический интерес. Но, во всяком случае, можно констатировать, что анализ дихотомических данных необходим во многих прикладных постановках.

### **3. Вероятностная модель дихотомических данных – люсиан**

Рассмотрим базовую вероятностную модель дихотомических данных – *бернуллиевский вектор* (в терминологии энциклопедии «Вероятность и математическая статистика» [8] – *люсиан*), т.е. конечную последовательность  $X = (X_1, X_2, \dots, X_k)$  независимых испытаний Бернулли  $X_i$  (т.е. случайных величин, принимающих два значения, а именно, 1 и 0),

для которых  $P(X_i = 1) = p_i$  и  $P(X_i = 0) = 1 - p_i$ ,  $i = 1, 2, \dots, k$ , причем вероятности  $p_i$  могут быть различны.

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

В связи с оцениванием по статистическим данным функции принадлежности нечеткой толерантности в 1970-е годы нами была построена теория случайных толерантностей с такими независимыми  $\delta_{ij}$ , что вероятности  $P(\delta_{ij} = 1) = p_{ij}$  произвольны [2]. 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 «случайное множество с независимыми элементами».

Был выявлен ряд областей, в которых полезен математический аппарат решения различных статистических задач, связанных с бернуллиевскими векторами (люсианами). Перечислим некоторые из них, 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 ММРІ (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, в медицине и в других ситуациях), и т.д.

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

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

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

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. [2, 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.

#### **4. Асимптотика растущей размерности и проверяемые гипотезы**

Продолжим изучение люсианов в рамках модели порождения дихотомических данных. Пусть  $A_1, A_2, \dots, A_s$  - независимые (между собой) люсианы с векторами параметров  $P_1, P_2, \dots, P_s$  соответственно. Будем использовать асимптотику  $s = \text{const}, k \rightarrow \infty$ . При этом число неизвестных параметров растет пропорционально объему данных. (Модели, в которых число неизвестных параметров фиксировано, рассмотрены в [9].)

В последние десятилетия (с начала 1970-х годов) в прикладной статистике все большее распространение получают постановки, в которых число неизвестных параметров растет вместе с объемом выборки. Результаты, полученные в подобных постановках, называют найденными

«в асимптотике растущей размерности» или «в асимптотике А.Н. Колмогорова» [10], перенося терминологию исследований по дискриминантному анализу на общий случай. Как известно, в задаче дискриминации [11] в две совокупности (т.е. отнесения вновь появляющегося объекта к одному из двух классов) академик АН СССР А.Н. Колмогоров (1903 - 1987) предложил рассматривать асимптотику

$$A \rightarrow \infty, N_i \rightarrow \infty, \frac{A}{N_i} \rightarrow \lambda_i > 0, i = 1, 2,$$

где  $A$  - размерность пространства (число признаков),  $N_i$  - объемы обучающих выборок,  $\lambda_i$  - константы,  $i = 1, 2$ . Эта асимптотика естественна при обработке многих видов технических, организационно-экономических, социологических, медицинских данных, поскольку число признаков, определяемых для каждого изучаемого объекта, респондента или пациента, обычно имеет тот же порядок, что и объем выборки.

Рассмотрим независимые (между собой) люсианы  $A_1, A_2, \dots, A_s$  с векторами параметров  $P_1, P_2, \dots, P_s$  соответственно. *Гипотезой согласованности* будем называть гипотезу

$$P_1 = P_2 = \dots = P_s. \quad (1)$$

В отличие от толерантностей, описываемых люсианами, для случайных ранжировок и разбиений под согласованностью понимают более частную гипотезу, предполагающую отрицание равномерности распределений (т.е. одинаковой вероятности появления каждой возможной ранжировки или разбиения), что соответствует замене проверки гипотезы (1) на проверку гипотезы

$$P_1 = P_2 = \dots = P_s = (1/2, 1/2, \dots, 1/2). \quad (2)$$

Как разъяснено в [2, 12], гипотеза (1) более адекватна конкретным задачам обработки реальных данных, например, экспертных оценок, чем (2). Поэтому полученные от экспертов данные, содержащие противоречия, целесообразно рассматривать как люсианы и проверять гипотезу (1), а не

подбирать ближайшие ранжировки или разбиения, после чего проверять согласованность методами теории случайных ранжировок или разбиений, как иногда рекомендуется.

Пусть  $A_1, A_2, \dots, A_m$  и  $B_1, B_2, \dots, B_n$  - независимые в совокупности люсианы длины  $k$ , одинаково распределенные в каждой группе с параметрами  $P(A)$  и  $P(B)$  соответственно. *Гипотезой однородности* называется гипотеза

$$P(A) = P(B).$$

В асимптотике растущей размерности принимаем, что объемы групп  $m$  и  $n$  постоянны, а размерность бернуллиевского вектора (число парных сравнений)  $k \rightarrow \infty$ .

Пусть  $(A_i, B_i)$ ,  $i = 1, 2, \dots, s$  - последовательность (фиксированной длины) пар люсианов. Пары предполагаются независимыми между собой. Требуется проверить гипотезу независимости  $A_i$  и  $B_i$ , т.е. внутри пар. В ранее введенных обозначениях *гипотеза независимости* - это гипотеза

$$P(X_{ij}(A) = 1, X_{ij}(B) = 1) = P(X_{ij}(A) = 1)P(X_{ij}(B) = 1), \\ i = 1, 2, \dots, s; j = 1, 2, \dots, k,$$

проверяемая в предположении

$$P_1(A) = P_2(A) = \dots = P_s(A), P_1(B) = P_2(B) = \dots = P_s(B).$$

В настоящем разделе излагается метод проверки гипотез о люсианах в асимптотике растущей размерности на примере гипотезы согласованности. Эти результаты сформулированы в [2, 12, 13]. Дальнейшее изучение проведено Г.В. Рыдановой, Т.Н. Дылько, Г.В. Раушенбахом, О.В. Филипповым, А.М. Никифоровым и др. Гипотеза однородности рассмотрена, например, в [13]. Методы проверки гипотезы однородности люсианов развиты и изучены Г.В. Рыдановой [14] на основе описанного ниже подхода. Она помимо доказательства предельных теорем провела подробное изучение скорости сходимости методом статистических испытаний.

Методы проверки согласованности люсианов нашли практическое применение, в частности, в медицине. Они были использованы в кардиологии при анализе данных кинетотопографии [13, 15, 16]. Эти методы включены в методические рекомендации Академии медицинских наук СССР и Ученого Медицинского Совета Минздрава СССР по управлению научными медицинскими исследованиями [17].

### **5. Метод проверки гипотез о люсианах в асимптотике растущей размерности**

Будем использовать дальнейшее развитие разработанного нами метода проверки гипотез по совокупности малых выборок (см., например, [9, с. 323–336]). Почему нельзя использовать иные подходы, имеющиеся в математической статистике, например, предназначенные для исключения влияния мешающих параметров? Поскольку число неизвестных параметров растет вместе с объемом выборки и пропорционально ему, эти параметры не являются мешающими (в том смысле, как этот термин понимается в теории математической статистики). Отметим, что согласно [18] равномерно наиболее мощных критериев не существует, поскольку параметров много. Не останавливаясь на других подходах математической статистики, констатируем необходимость применения метода проверки гипотез по совокупности малых выборок.

Пусть имеются  $k$  выборок, независимых между собой. Пусть при справедливости нулевой гипотезы по каждой из выборок можно построить несмещенную оценку  $\xi_i \in R^p$  векторного нуля  $0 \in R^p$ , где  $p \geq 1$ ,  $i = 1, 2, \dots, k$ . Другими словами, пусть распределение  $i$ -ой выборки описывается параметром  $\theta_i$ , лежащим в произвольном пространстве, а нулевая гипотеза, очевидно, состоит в том, что  $\theta_i \in \Theta_{0i}$ , где  $\Theta_{0i}$  - собственное подмножество множества  $\{\theta_i\}$ . Предполагается, что можно по  $i$ -ой выборке вычислить статистику  $\xi_i$  такую, что

$$M\xi_i = 0 \quad (3)$$

при всех  $\theta_i \in \Theta_{0i}$ . Очевидно,  $\xi_i \equiv 0$  удовлетворяют (1). Однако для рассматриваемого метода необходимо, чтобы при всех  $\theta_i \in \Theta_{0i}$  ковариационная матрица вектора  $\xi_i$  была ненулевой:

$$Cov(\xi_i) = M(\xi_i^T \xi_i) \neq 0. \quad (4)$$

В теории математической статистики иногда используют понятие полноты параметрического семейства распределений. Если рассматриваемое семейство является полным - а так и есть для люсианов, - то не существует достаточной статистики, удовлетворяющей одновременно условиям (1) и (2) (см., например, [19, §§2.12 - 2.14]). Поэтому будем использовать статистики, не являющиеся достаточными.

Следующее предположение - ковариационные матрицы статистик  $\xi_i$ , т.е.  $Cov(\xi_i)$ , также допускают несмещенные оценки  $S_i$  по тем же выборкам:

$$M(S_i) = Cov(\xi_i) \quad (5)$$

при всех  $\theta_i \in \Theta_{0i}$ .

Рассматриваемый метод основан на том, что поскольку случайные вектора  $\xi_i$  определяются по независимым между собой выборкам, то  $\xi_i$  независимы в совокупности, а потому случайный вектор

$$\xi = \sum_{i=1}^k \xi_i \quad (6)$$

является суммой независимых случайных векторов, имеет в силу (3) нулевое математическое ожидание, а его ковариационная матрица равна

$$C_k = \sum_{i=1}^k Cov(\xi_i).$$

При справедливости многомерной центральной предельной теоремы (простейшее условие справедливости этой теоремы для  $\xi_i$  в случае люсианов - отделенность от 0 и 1 всех элементов матриц  $P_j$ , равномерная по  $s$  и  $k$ ) вектор  $\xi$  является асимптотически нормальным, т.е. при  $k \rightarrow \infty$  распределение  $\xi$  сближается (в смысле, раскрытом в [9, Приложение 1]) с

многомерным нормальным распределением  $N(0; C_k)$ .

Однако эту сходимость нельзя непосредственно использовать для проверки исходной гипотезы, поскольку матрица  $C_k$  неизвестна статистику. Необходимо оценить эту матрицу по статистическим данным. В силу (5) в качестве оценки  $C_k$  естественно использовать

$$C_k^* = \sum_{i=1}^k S_i.$$

Простейшая формулировка условий справедливости такой замены - предположение о том, что к последовательности  $S_i$  можно применить закон больших чисел. А именно, пусть существует неотрицательно определенная матрица  $C$  такая, что при  $k \rightarrow \infty$

$$\frac{1}{k}(C_k^* - C_k) \rightarrow 0, \quad \frac{1}{k}C_k \rightarrow C. \quad (7)$$

В силу результатов приложения 1 из асимптотической нормальности  $\xi$  и соотношений (7) следует, что распределение статистики

$$\eta = \frac{1}{\sqrt{k}}\xi$$

сходится к нормальному распределению  $N(0; C)$ . При этом, если некоторый случайный вектор  $\tau$  имеет распределение  $N(0; C)$ , то распределение случайной величины  $q(\eta)$  сходится к распределению  $q(\tau)$  для произвольной интегрируемой по Риману по любому кубу функции  $q: R^p \rightarrow R^1$  [9, Приложение 1]. Для проверки нулевой гипотезы предлагается пользоваться статистикой  $q(\eta)$  при подходящей функции  $q$ , а процентные точки брать соответственно распределению  $q(\tau)$ . В этом и состоит рассматриваемый метод проверки гипотез о люсианах в асимптотике растущей размерности. Для реальных расчетов целесообразно использовать линейные или квадратические функции  $q$  от координат вектора  $\eta$ .

Отклонения от нулевой гипотезы приводят, как правило, к нарушению равенств (3) и (4). Случайный вектор  $\eta$  при этом обычно

остаётся асимптотически нормальным, но с другими параметрами, что может быть обычным образом использовано для построения оптимального решающего правила, соответствующего заданной альтернативе (например, согласно лемме Неймана-Пирсона). Поведение при альтернативах для некоторых гипотез изучено в [13, 14], здесь его не будем рассматривать, поскольку вычисление мощности не требует новых идей.

## 6. Несмещенные оценки параметров асимптотического распределения вектора попарных расстояний

Применим описанный выше метод для проверки гипотезы согласованности люсианов. Исходные данные - люсианы

$$A_j = (X_{1j}, X_{2j}, \dots, X_{kj}), j = 1, 2, \dots, s.$$

В качестве  $i$ -й выборки возьмем совокупность испытаний Бернулли, стоящих на  $i$ -м месте в рассматриваемых люсианах:

$$X_{i1}, X_{i2}, \dots, X_{is}. \quad (8)$$

При справедливости нулевой гипотезы в (8) стоят независимые испытания Бернулли с одной и той же вероятностью успеха  $p_i$ ; при нарушении нулевой гипотезы согласованности независимость испытаний Бернулли сохраняется, но вероятности успеха могут различаться.

В качестве вектора  $\xi$ , на основе которого строятся статистики для проверки согласованности, будем использовать вектор попарных расстояний между люсианами

$$\xi = \{d(A_p, A_q), 1 \leq p < q \leq s\}, \quad (9)$$

в котором пары  $(p, q)$  упорядочены лексикографически,

$$d(A_p, A_q) = \sum_{i=1}^k \mu_i |X_{ip} - X_{iq}|, \quad \mu_i > 0. \quad (10)$$

В [9, гл. 1] это расстояние выведено из некоторой системы аксиом (напомним, что совокупность векторов из 0 и 1 размерности  $k$  находится во взаимно-однозначном соответствии с совокупностью подмножеств

множества из  $k$  элементов; при этом 1 соответствует тому, что элемент входит в подмножество, а 0 - что не входит).

Из вида расстояния в формуле (10) следует, что введенный в (9) вектор  $\xi$  имеет вид (6) с

$$\xi_i = \mu_i \{ |X_{ip} - X_{iq}|, 1 \leq p < q \leq s \}. \quad (11)$$

Следовательно, для применения описанного выше метода проверки гипотез о люсианах в асимптотике растущей размерности достаточно построить на основе вектора  $\xi_i$  из (11) несмещенную оценку  $\theta$  и найти несмещенную оценку ковариационной матрицы этой оценки.

Чтобы применить общую схему, необходимо начать с построения статистики  $\beta$  такой, чтобы при всех  $p_i$  имело место равенство

$$M(|X_{ip} - X_{iq}| - \beta) = 0, \quad 1 \leq p < q \leq s.$$

Элементарный расчет дает:

$$M|X_{ip} - X_{iq}| = 2p_i(1 - p_i).$$

Как известно [4, с.56-57], несмещенная оценка многочлена

$$f(p) = \sum_{h=0}^m a_h p^h$$

по результатам  $m$  независимых испытаний Бернулли с вероятностью успеха  $p$  в каждом имеет вид

$$f^*(p) = \sum_{h=0}^m a_h \frac{\gamma^{[h]}}{m^{[h]}}, \quad (12)$$

где  $\gamma$  - общее число успехов в  $m$  испытаниях и использовано обозначение

$$n^{[h]} = n(n - 1) \dots (n - h + 1).$$

Ясно, что многочлены степени  $m + 1$  и более высокой невозможно несмещенно оценить по результатам  $m$  испытаний.

В случае  $f(p) = 2p(1 - p)$  в соответствии с (12) получаем несмещенную оценку

$$\beta = \frac{2}{m-1} \left( \gamma - \frac{\gamma^2}{m} \right). \quad (13)$$

Таким образом, можно применять общий метод проверки гипотез о люсианах в асимптотике растущей размерности  $s$

$$\xi_i = \mu_i (\{|X_{ip} - X_{iq}|, 1 \leq p < q \leq s\} - \beta_i e),$$

где коэффициенты  $\beta_i$  определяются с помощью формулы (13) по  $\gamma_i$  - общему числу единиц, стоящих на  $i$ -м месте в люсианах  $A_1, A_2, \dots, A_s$ , а  $e$  - вектор размерности  $s(s - 1)/2$  с единичными координатами. Тогда несмещенная оценка  $0$ , о которой идет речь в методе проверки гипотез по совокупности малых выборок, имеет вид

$$\xi = \{d(A_p, A_q), 1 \leq p < q \leq s\} - \sum_{i=1}^k \mu_i \beta_i e.$$

Для использования статистики типа  $\eta$ , распределение которой приближается с помощью нормального распределения

$$N\left(0; \frac{1}{k} \sum_{i=1}^k S_i\right),$$

необходимо уметь несмещенно оценивать ковариационные матрицы  $Cov(\xi_i)$ . Для этого достаточно найти математические ожидания элементов матрицы  $M(\xi_i^T \xi_i)$  как функции (многочлены) от  $p_i$ , а затем использовать формулу (12) для получения несмещенных оценок.

Вычисление матрицы  $M(\xi_i^T \xi_i)$  хотя и трудоемко, но не содержит каких-либо принципиальных трудностей. В [13] вычислены диагональные элементы рассматриваемой матрицы. Вычисление занимает около 2,5 книжных страниц (с.299 - 301). Поэтому здесь приведен только окончательный итог.

Обозначим для краткости  $p_i = p$ . В [13] показано, что

$$D = D(|X_{ip} - X_{iq}| - \beta_i) = \left(2 - \frac{4}{s}\right) p(1-p) - 4 \frac{(s-2)(s-3)}{s(s-1)} p^2(1-p)^2.$$

Если двухэлементные множества  $\{p, q\}$  и  $\{r, t\}$  не имеют ни одного общего элемента, то

$$C_1 = M(|X_{ip} - X_{iq}| - \beta_i)(|X_{ir} - X_{it}| - \beta_i) = -\frac{4}{s} p(1-p) + \frac{8(2s-3)}{s(s-1)} p^2(1-p)^2,$$

а если имеют ровно один общий элемент, то

$$C_2 = M(|X_{ip} - X_{iq}| - \beta_i)(|X_{ir} - X_{it}| - \beta_i) = \left(1 - \frac{4}{s}\right)p(1-p) - 4 \frac{(s-2)(s-3)}{s(s-1)} p^2(1-p)^2.$$

С помощью формулы (12) получаем несмещенные оценки для  $D$ ,  $C_1$  и  $C_2$  как многочленов от  $p$ :

$$D^* = \frac{2\gamma_i(s-\gamma_i)}{s^2(s-1)^2} \{(s-2)(s-1) - 2(\gamma_i-1)(s-\gamma_i-1)\},$$

$$C_1^* = \frac{4\gamma_i(s-\gamma_i)}{s^2(s-1)} \left\{ \frac{2(2s-3)(\gamma_i-1)(s-\gamma_i-1)}{(s-1)(s-2)(s-3)} - 1 \right\},$$

$$C_2^* = \frac{\gamma_i(s-\gamma_i)}{s^2(s-1)^2} \{(s-4)(s-1) - 4(\gamma_i-1)(s-\gamma_i-1)\}.$$

С помощью трех чисел  $D^*, C_1^*, C_2^*$  выписывается несмещенная оценка матрицы ковариаций вектора  $\xi_i/\mu_i$ , которую обозначим  $B_i$ . Тогда асимптотически нормальный вектор  $\xi$  имеет нулевое математическое ожидание и ковариационную матрицу, несмещенно и состоятельно (в смысле соотношений (7)) оцениваемую с помощью

$$Cov(\xi)^* = \sum_{i=1}^k \mu_i^2 B_i. \quad (14)$$

Асимптотическая нормальность доказывается, естественно, в схеме серий. Достаточным условием является существование положительной константы  $\varepsilon$  такой, что

$$\mu_i \geq \varepsilon, \quad \frac{1}{\mu_i} \geq \varepsilon, \quad p_i \geq \varepsilon, \quad 1-p_i \geq \varepsilon \quad (15)$$

при всех  $k$  и  $i$ ,  $1 \leq i \leq k$ .

Поскольку  $D$ ,  $C_1$  и  $C_2$  являются многочленами четвертой степени от  $p$ , то несмещенные оценки для них существуют при  $s \geq 4$ . Если же  $s < 4$ , то несмещенных оценок не существует. Поэтому указанным методом проверять согласованность можно лишь при числе люсианов  $s \geq 4$ .

## 7. Проверка согласованности люсианов

Пусть  $\alpha$  - нормально распределенный случайный вектор размерности  $s(s - 1)/2$  с нулевым математическим ожиданием и ковариационной матрицей, определенной формулой (14). Согласно результатам, приведенным в [9, Приложение 1], для любой действительнзначной функции  $f$ , интегрируемой по Риману по любому гиперкубу, распределения случайных величин  $f(\xi)$  и  $f(\alpha)$  сближаются при  $k \rightarrow \infty$ . Это означает, что вместо распределения случайной величины  $f(\xi)$  для построения критериев проверки гипотез можно использовать распределение случайной величины  $f(\alpha)$ . Более того, аналогичный результат верен при замене  $f$  на  $f_n$  (при слабых внутриматематических условиях регулярности, наложенных на последовательность функций  $f_n$ ). Следовательно, для проверки гипотезы согласованности люсианов можно пользоваться любой статистикой  $f_n(\xi)$ , для которой могут быть вычислены с помощью соответствующего программного продукта или заранее табулированы процентные точки распределения  $f_n(\alpha)$ , аппроксимирующего распределение  $f_n(\xi)$ .

В частности, можно использовать линейные статистики, представляющие собой скалярное произведение случайного вектора  $\xi$  и некоторого заданного детерминированного вектора коэффициентов  $a$ , т.е.

$$(\xi, a) = \sum_{i=1}^k \left( \mu_i \sum_{1 \leq j < t \leq s} a_{jt} (|X_{ij} - X_{it}| - \beta_i) \right). \quad (16)$$

Линейные статистики имеют нулевое математическое ожидание и дисперсию, очевидным образом выражающуюся через матрицу коэффициентов  $\|a_{ij}\|$  и числа  $D$ ,  $C_1$  и  $C_2$ , а потому несмещенно и состоятельно оцениваемую с помощью с помощью выписанных выше оценок для  $D$ ,  $C_1$  и  $C_2$ .

Отметим, что  $(\xi, a) = 0$  при  $a_{ij} \equiv 1$ ,  $1 \leq j < t \leq s$ . Это следует как из непосредственного вычисления дисперсии  $(\xi, a)$ , так и из того, что  $(\xi, a)$  в

рассматриваемом случае выражается через достаточную статистику  $(\gamma_1, \gamma_2, \dots, \gamma_k)$  и является несмещенной оценкой нуля, а семейство биномиальных распределений полно, т.е. существует только одна несмещенная оценка нуля - тождественный нуль. Таким образом, сумма координат вектора  $\xi$ , т.е. непосредственный аналог коэффициента ранговой конкордации Кендалла - Смита из теории ранговой корреляции [20, 21], тождественно равна 0.

Распределение статистики (16) при альтернативах изучено в работе [14].

Рассмотрим два частных случая.

*Первый частный случай.* Проверка согласованности двух определенных люсианов (ответов двух экспертов),  $j$ -го и  $t$ -го, может осуществляться с помощью статистики (16), в которой отличен от 0 только член с  $a_{jt} = 1$ . Оценкой дисперсии является  $D^*$ .

*Второй частный случай.* Пусть необходимо проверить согласованность люсианов с одним из них, скажем, с  $j$ -м (например, люсианы отражают мнения экспертов, а  $j$ -й из них является наиболее компетентным - по априорной оценке, или «лицом, принимающим решения», или его мнение сильно отличается от мнений остальных). Это можно сделать с помощью статистики (16), в которой

$$a_{jt} = 1, t = j + 1, j + 2, \dots, s; \quad a_{tj} = 1, t = 1, 2, \dots, j - 1; \\ a_{qt} = 0, q \neq j, t \neq j, 1 \leq q < t \leq s.$$

Другими словами, она имеет вид

$$W = \sum_{t=1}^s d(A_j, A_t) - (s-1) \sum_{i=1}^k \mu_i \beta_i, \quad (17)$$

где расстояние  $d$  между люсианами определено в (10), а  $\beta_i$  - в (13) с заменой  $m$  на  $s$  и  $\gamma$  на  $\gamma_i$ . Используя полученные ранее несмещенные оценки элементов ковариационной матрицы, нетрудно показать, что несмещенная и состоятельная (в смысле формулы (7) выше) оценка дисперсии  $W$  имеет

вид

$$D^*(W) = \sum_{i=1}^k \mu_i^2 \frac{\gamma_i(s - \gamma_i)}{s^2} \{(s - 2)^2 - 4(\gamma_i - 1)(s - \gamma_i - 1)\}.$$

Тогда при выполнении некоторых внутриматематических условий регулярности, например, условий (15), распределение статистики

$$\frac{1}{\sqrt{D^*(W)}} W$$

сходится при  $k \rightarrow \infty$ ,  $s = \text{const}$  к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1 (при справедливости гипотезы (1) согласованности люсианов).

Статистика (17) наряду со статистикой, предназначенной для проверки гипотезы однородности люсианов, включена в «Методические рекомендации» АМН СССР и УМС Минздрава СССР [17]. Последнюю статистику не расписываем здесь, поскольку для этого не требуются новые идеи.

## 8. Различные подходы к понятию согласованности

Обсудим условия, при выполнении которых люсианы естественно считать согласованными (а экспертов, чьи мнения отражают люсианы, имеющими единое мнение, искаженное случайными ошибками), т.е. обсудим различные методы проверки гипотезы (1).

*Полное индивидуальное согласие* имеет место, если никакие два эксперта не являются «несогласованными». Уровень значимости определяется описанным выше способом (первый частный случай). Однако наличие одной или нескольких пар экспертов, чьи мнения нельзя считать согласованными, не свидетельствует о необходимости отклонения гипотезы (1), поскольку парных проверок проводится много, а именно,  $s(s - 1) \geq 6$ , а способы установления уровня значимости при множественных проверках, зависящих между собой, к настоящему времени плохо

разработаны [22]. Проблема множественных проверок для количественных признаков обсуждается А.А. Любищевым [23, с.36-39], выход дается дисперсионным анализом. Можно брать не все попарные проверки, а только для  $[s/2]$  пар люсианов, причем разбиение на пары проводить независимо от принятых люсианами значений, как это делает Т.Н. Дылько [24]. Тогда для проверки гипотезы (1) на уровне значимости  $\alpha$  надо брать для проверки в каждой паре уровень значимости  $\beta$ , где  $\beta$  рассчитывается понятным образом, приближенно  $\beta = \alpha / [s/2]$ .

*Полное согласие в целом* означает, что для любого эксперта мнения всех остальных оказываются с ним согласованными при использовании статистики (17) (второй частный случай). Отсутствие подобного согласия для одного или нескольких экспертов не означает отклонения гипотезы согласованности люсианов (1) - по тем же причинам, что и в предыдущем случае.

*Минимальное согласие* имеют мнения экспертов, когда хотя бы для одного из них гипотеза согласованности не отвергается с помощью статистики (17). В этом случае групповое мнение целесообразно строить, выделяя «ядро», о чем подробнее сказано ниже.

Расстояние  $d$  между люсианами (см. формулу (10)) введено аксиоматически в [9, гл. 1] (напомним, что реализацию люсиана можно рассматривать как подмножество конечного множества). Там же из иной системы аксиом выведено другое расстояние -  $D$ -метрика. Рассмотрим проверку согласованности люсианов с использованием  $D$ -метрики. В этом случае расстояние между люсианами  $A_1$  и  $A_2$  имеет вид

$$D(A_1, A_2) = \begin{cases} \frac{d(A_1, A_2)}{T(A_1, A_2)}, & T(A_1, A_2) \neq 0, \\ 0, & T(A_1, A_2) = 0, \end{cases}$$

где

$$T(A_1, A_2) = \sum_{i=1}^k \mu_i \max(X_{i1}, X_{i2}).$$

Ясно, что теория, основанная на  $D$ -метрике, из-за наличия знаменателя в только что приведенной формуле существенно сложнее теории, основанной на метрике  $d$ . Ясно, что описанный выше метод проверки гипотез о люсианах в асимптотике растущей размерности применить не удастся. Чтобы продемонстрировать существенное усложнение ситуации, опишем лишь асимптотическое поведение расстояния  $D(A_1, A_2)$  между двумя люсианами, анонсированное в [25].

*Теорема.* Пусть  $p_{1i}$  и  $p_{2i}$  отделены от 0 и 1, а  $\mu_i$  отделены от 0 и  $+\infty$ . Тогда расстояние  $D(A_1, A_2)$  между люсианами  $A_1$  и  $A_2$  асимптотически нормально при  $k \rightarrow \infty$  с параметрами

$$t_k = \frac{N_1}{N_2}, \quad q_k = \frac{N_1}{N_2} \sqrt{\frac{N_3}{N_1^2} + \frac{N_4}{N_2^2} - 2 \frac{N_5}{N_1 N_2}},$$

т.е. для любого числа  $x$  справедливо предельное соотношение

$$\lim_{k \rightarrow \infty} P \left\{ \frac{D(A_1, A_2) - t_k}{q_k} \leq x \right\} = \Phi(x),$$

где  $\Phi(x)$  - функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1.

Величины  $N_j, j = 1, 2, 2, 4, 5$ , выражаются через  $\mu_i$  и величины

$$p_{3i} = p_{1i} + p_{2i} - 2p_{1i} p_{2i}, \quad p_{4i} = p_{1i} + p_{2i} - p_{1i} p_{2i}$$

следующим образом:

$$N_1 = \sum_{i=1}^k \mu_i p_{3i}, \quad N_2 = \sum_{i=1}^k \mu_i p_{4i}, \quad N_3 = \sum_{i=1}^k \mu_i^2 p_{3i} (1 - p_{3i}),$$

$$N_4 = \sum_{i=1}^k \mu_i^2 p_{4i} (1 - p_{4i}), \quad N_5 = \sum_{i=1}^k \mu_i^2 p_{3i} (1 - p_{4i}).$$

*Следствие 1.* Пусть  $p_{1i} = p_1$  и  $p_{2i} = p_2$  при всех  $i, k$ , причем  $p_1$  и  $p_2$  лежат внутри отрезка  $(0; 1)$ . Пусть  $\mu_i$  отделены от 0 и  $+\infty$ . Тогда расстояние  $D(A_1, A_2)$  между люсианами  $A_1$  и  $A_2$  асимптотически нормально при  $k \rightarrow \infty$  с параметрами

$$t_k = \frac{p_3}{p_4}, \quad q_k^2 = \frac{p_1 p_2 p_3}{p_4^3} \frac{\sum_{i=1}^k \mu_i^2}{\left(\sum_{i=1}^k \mu_i\right)^2},$$

где

$$p_3 = p_1 + p_2 - 2p_1 p_2, \quad p_4 = p_1 + p_2 - p_1 p_2.$$

*Следствие 2.* Пусть в предположениях следствия 1  $p_1 = p_2 = p$  и  $\mu_i = 1$  при всех  $i, k$ . Тогда

$$t_k = \frac{2(1-p)}{2-p}, \quad q_k = \frac{2(1-p)}{k(2-p)^3}.$$

*Замечание.* Пусть вследствие 2  $p = 1/2$ . Тогда  $A_1$  и  $A_2$  - люсианы, равномерно распределенные на множестве всех последовательностей из 0 и 1 длины  $k$ . В частности, эти люсианы могут соответствовать независимым случайным множествам, равномерно распределенным на совокупности всех подмножеств конечного множества из  $k$  элементов, или независимым толерантностям, равномерно распределенным на множестве всех толерантностей, определенных на множества из  $m$  элементов, где  $m(m-1)/2 = k$ . По следствию 2 расстояние между люсианами  $D(A_1, A_2)$  асимптотически нормально с математическим ожиданием 0,667 и дисперсией  $0,296 k^{-1}$ . Напомним, что распределения коэффициентов ранговой корреляции Кендалла и Спирмена изучены (в основном) лишь при условии равномерности распределения случайных ранжировок на множестве всех возможных ранжировок фиксированного числа объектов. Для теории люсианов случай равномерности распределения - весьма частный, а для теории ранжировок - основной. Как уже говорилось, отказ от равномерности - привлекательная черта теории люсианов.

## 9. Классификация люсианов

Отсутствие согласованности в одном из перечисленных выше смыслов позволяет сделать заключение о целесообразности разбиения всех

люсианов (например, если они выражают мнения экспертов) на группы близких между собой, т.е. о целесообразности классификации люсианов, точнее, их кластер-анализа. Поскольку введена мера близости между люсианами  $d(A_1, A_2)$  или  $D(A_1, A_2)$ , то естественным является следующий способ действий: провести разбиение на кластеры с помощью одного из алгоритмов, основанных на использовании меры близости, а затем проверить мнения в каждом классе на согласованность. Однако применение того или иного алгоритма кластер-анализа, вообще говоря, может нарушить предпосылки описанных выше способов описанных выше способов проверки согласованности (ср. обсуждение похожей проблемы, связанной с применением регрессионного анализа после кластер-анализа, в [26, гл.11]). Поэтому опишем методы классификации, опирающиеся на результаты проверки согласованности.

Разбиение на кластеры, внутри каждого из которых имеет место «полное индивидуальное согласие», может быть проведено с помощью агломеративного иерархического алгоритма «дальнего соседа», дополненного ограничением сверху на диаметр кластера. Это ограничение строится из статистических соображений, в отличие от методов, обычно используемых в кластер-анализе [11]. При этом в качестве меры близости между люсианами используют не расстояния  $d$  или  $D$ , а модуль статистики, применяемой для проверки согласованности двух люсианов, т.е. статистики (16), в которой только одно из чисел  $a_{ij}$  отлично от 0. Упомянутое ограничение таково: диаметр кластера не должен превосходить процентной точки предельного распределения, соответствующей используемому при анализе рассматриваемых данных уровню значимости (можно порекомендовать 5%-й уровень значимости). В результате работы алгоритма получим кластеры, в которых имеется «полное индивидуальное согласие», причем объединение любых двух кластеров приведет к исчезновению этого свойства у объединения.

Поскольку способ выделения итогового разбиения из иерархического дерева разбиений имеет вероятностно-статистическое обоснование, изложенное выше, то описанный метод классификации люсианов следует считать - в терминологии [27] - не методом анализа данных, а вероятностно-статистическим методом.

Кластеры «с полным согласием в целом» могут быть получены с помощью агломеративного иерархического алгоритма, в котором мерой близости двух кластеров является максимальное значение модуля статистики (17), когда  $j$  пробегает номера мнений (люсианов), вошедших в объединение рассматриваемых кластеров, а суммирование в (17) проводится по всем люсианам в этом объединении. Ограничение сверху на меру близости кластеров определяется процентной точкой предельного распределения статистики  $W$ , заданной формулой (17).

Кластеры «с минимальным согласием» можно получить, при фиксированном  $j$  выделяя совокупность люсианов, согласованных с  $A_j$  в смысле статистики  $W$  из (17).

На основе двух рассмотренных выше частных случаев линейной статистики (16) можно строить и другие способы классификации. Например, для каждого люсиана  $A_m$  можно выделить кластер «типа шара» (см. [26, гл.5]) из люсианов, попарно согласованных с  $A_m$ . Все такие способы имеют вероятностно-статистическое обоснование, и потому к ним относится сказанное выше относительно выделения кластеров «с полным индивидуальным согласием».

*Замечание.* Проверка согласованности приведенными выше критериями может привести к отрицательному результату двумя способами - либо значение статистики окажется слишком большим, либо слишком малым. Первое означает, что гипотеза согласованности люсианов (1) неверна, вторая - что неверна вероятностная модель реального явления или процесса, основанная на люсианах. С необходимостью учета второй

возможности мы столкнулись при применении теории люсианов для анализа данных топокарт, полученных при проведении кинетокардиографии у больных инфарктом миокарда [15, 16].

### 10. Нахождение среднего

В результате классификации получаем согласованные (в одном из указанных выше смыслов) группы люсианов. Для каждой из них полезно рассмотреть среднее. В зависимости от конкретных приложений в прикладных исследованиях применяют либо среднее в виде последовательностей 0 и 1, т.е. в виде реализации люсиана, либо среднее в виде последовательности оценок вероятностей  $(p_1, p_2, \dots, p_k)$ . Кроме того, оно может находиться либо с помощью методов, подавляющих «засорения» («выбросы»), либо без учета возможности засорения. Рассмотрим все четыре возможности.

В соответствии с подходом главы 2 при отсутствии засорения эмпирическое среднее ищется как решение задачи

$$\sum_{j=1}^m d(A_j, A) \rightarrow \min_{A \in X}, \quad (18)$$

где  $A_1, A_2, \dots, A_m$  - люсианы, входящие в рассматриваемый кластер,  $X$  - множество, которому принадлежит среднее. Если  $X$  - совокупность последовательностей из 0 и 1, то правило (18) дает решение по правилу большинства.

Если  $X$  - пространство последовательностей вероятностей, то решением задачи (18) является та же последовательность 0 и 1, что и в первом случае. Поэтому в качестве среднего вместо решения задачи (18) целесообразно рассматривать просто последовательность частот.

Асимптотическое поведение средних при  $m \rightarrow \infty$  вытекает из законов больших чисел [28], теорем, описывающих асимптотику решений экстремальных статистических задач [29], и теоремы Муавра-Лапласа

соответственно.

В работе [30] при анализе результатов эксперимента показано, что ответы реальных экспертов разбиваются на многочисленное «ядро», расположенное вокруг истинного мнения, и отдельных «диссидентов», разбросанных по периферии. Причем оценка истинного мнения по «ядру» является более точной, чем по всей совокупности, поскольку мнения «диссидентов» не отражают истинного мнения. Поэтому для построения группового мнения, в том числе среднего для совокупности люсианов, отражающих мнения экспертов, естественно применять методы, подавляющие мнения «диссидентов», что соответствует методологии робастности.

«Ядро» может быть построено следующим образом. Решается задача (18) с конечным множеством  $X$ , состоящим из всех исходных люсианов:  $X = \{A_1, A_2, \dots, A_m\}$ , т.е. из результатов наблюдений выбирается тот, что находится «в центре» совокупности результатов наблюдений. Пусть  $A_j$  является решением этой задачи. В качестве ядра предлагается рассматривать совокупность всех люсианов, которые попарно согласованы с  $A_j$ . Другой вариант: рассматривается кластер с «полным внутренним согласием», куда входит  $A_j$ . (При этом, очевидно, должно быть изменено (уменьшено) критическое значение критерия по сравнению с процедурой, приведшей к выделению группы, нахождением группового мнения которой мы занимаемся.) Затем групповое мнение ищется лишь для элементов «ядра». Описанная процедура особенно необходима в случае, когда не было предварительного разбиения совокупности люсианов на группы, согласованные друг с другом. Новым по сравнению с [30] является придание вероятностного смысла порогу, выделяющему «ядро».

Обобщая идею выделения «ядра», приходим к «взвешенным итеративным методам оценивания среднего» (ВИМОП - оценкам среднего), введенным и первоначально изученным в работе [31]. Их

применение для люсианов не требует специальных рассмотрений.

Таким образом, в настоящей статье представлен ряд методов обработки специального вида объектов нечисловой природы - люсианов. При этом для решения одной и той же задачи, например, задачи классификации, предлагается ряд методов, точно так же, как для решения классической задачи проверки однородности двух независимых выборок имеется большое число методов [26, гл.4].

Теория люсианов – часть статистики объектов нечисловой природы [32]. Разработка этой теории ведется в соответствии с новой парадигмой математической статистики [33] и анализа статистических и экспертных данных в задачах экономики и управления [34]. Распределение люсиана описывается нечетким множеством, а потому рассматриваемая теория относится к системной нечеткой интервальной математике [35, 36].

### Литература

1. Суппес П., Зинес Дж. Основы теории измерений // Психологические измерения. - М.: Мир, 1967. - С.9 - 110.
2. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979. - 296 с.
3. Беляев Ю.К. Вероятностные методы выборочного контроля. - М.: Наука, 1975. - 408 с.
4. Лумельский Я.П. Статистические оценки результатов контроля качества. - М.: Изд-во стандартов, 1979. - 200 с.
5. Орлов А.И. Организационно-экономическое моделирование : учебник : в 3 ч. Ч.3. Статистические методы анализа данных. - М.: Изд-во МГТУ им. Н.Э. Баумана, 2012. - 624 с.
6. Дэвид Г. Метод парных сравнений. - М.: Статистика, 1978.- 144 с.
7. Орлов А.И. Организационно-экономическое моделирование: учеб. Ч.2. Экспертные оценки. - М.: Изд-во МГТУ им. Н.Э. Баумана, 2011. - 486 с.
8. Вероятность и математическая статистика: Энциклопедия / Гл. ред. Ю.В. Прохоров. - М.: Большая Российская энциклопедия, 1999. - 910 с.
9. 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.
10. Орлов А.И. Парные сравнения в асимптотике Колмогорова. – В сб.: Экспертные оценки в задачах управления. - М.: Изд-во Института проблем управления АН СССР, 1982. - С. 58-66.
11. 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>

12. Орлов А.И. Статистика объектов нечисловой природы и экспертные оценки. – В сб.: Экспертные оценки / Вопросы кибернетики. Вып.58. – М.: Научный Совет АН СССР по комплексной проблеме "Кибернетика", 1979. – С.17-33.

13. Орлов А.И. Случайные множества с независимыми элементами (люсианы) и их применения. – В сб.: Алгоритмическое и программное обеспечение прикладного статистического анализа. Ученые записки по статистике, т.36. – М.: Наука, 1980. – С. 287-308.

14. Рыданова Г.В. Некоторые вопросы статистического анализа случайных бинарных векторов. Дисс. ... канд. физ.-мат. наук. – М.: МГУ, ф-т вычислительной математики и кибернетики, 1987. – 139 с.

15. Аксенова Г.А., Кузьмина Е.С., Орлов А.И., Розова Н.К. Кинетотопография в диагностике инфаркта миокарда. – В сб.: Актуальные вопросы клинической и экспериментальной медицины. – М.: 4 Главное Управление при Минздраве СССР, 1979. С.24-26.

16. Попов В.Г., Аксенова Г.А., Орлов А.И., Розова Н.К., Кузьмина Е.С. Кинетокардиография в определении зон асинергии у больных инфарктом миокарда. – Журнал «Клиническая медицина». 1982. Т.LX. №.3. С.25-30.

17. Методические рекомендации по проведению экспертной оценки планируемых и законченных научных работ в области медицины (по проблемам союзного значения) / Составители: Г.В. Раушенбах, О.В. Филиппов. – М.: АМН СССР - Ученый медицинский совет Минздрава СССР, 1982. – 36 с.

18. Леман Э. Проверка статистических гипотез. – М.: Наука, 1979. – 408 с.

19. Боровков А.А. Математическая статистика / Учебное пособие для вузов. – М.: Наука, 1984. – 472 с.

20. Кендэл М. Ранговые корреляции. – М.: Статистика, 1975. – 216 с.

21. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. – М.: Наука, 1965 (1-е изд.), 1968 (2-е изд.), 1983 (3-е изд.).

22. Орлов А.И. Проблема множественных проверок статистических гипотез // Заводская лаборатория. Диагностика материалов. 1996. Т.62. № 5. С.51 - 54.

23. Любищев А.А. Дисперсионный анализ в биологии. – М.: Изд-во МГУ, 1986. – 200 с.

24. Дылько Т.Н. Проверка гипотез в экспертном оценивании // Вестник Белорусского государственного университета / Сер. 1: Физика, математика и механика. 1988. № 2. С.36 - 40.

25. Орлов А.И., Раушенбах Г.В. Метрика подобия: аксиоматическое введение, асимптотическая нормальность. – В сб.: Статистические методы оценивания и проверки гипотез. Межвузовский сборник научных трудов. – Пермь: Изд-во Пермского государственного университета, 1986, с.148-157.

26. Орлов А.И. Эконометрика. Учебник для вузов. Изд. 3-е, исправленное и дополненное. – М.: Изд-во «Экзамен», 2004. – 576 с.

27. Орлов А.И., Миронова Н.Г., Фомин В.Н., Черчинцев А.Н. Рекомендации. Прикладная статистика. Методы обработки данных. Основные требования и характеристики. – М.: ВНИИ стандартизации Госстандарта СССР, 1987. – 62 с.

28. 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>

29. Орлов А.И. Асимптотика решений экстремальных статистических задач // Анализ нечисловых данных в системных исследованиях. Сборник трудов. Вып.10. - М.: Всесоюзный научно-исследовательский институт системных исследований, 1982. С.4 - 12.

30. Тюрин Ю.Н., Василевич А.П. К проблеме обработки рядов ранжировок // Статистические методы анализа экспертных оценок. Ученые записки по статистике, т.29. - М.: Наука, 1977. - С.96 - 111.

31. Орлов А.И. Некоторые вероятностные вопросы теории классификации // Прикладная статистика. Ученые записки по статистике, т.45. - М.: Наука, 1983. - С.166-179.

32. 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>

33. 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. 187 – 213. – IDA [article ID]: 0901306013. – Режим доступа: <http://ej.kubagro.ru/2013/06/pdf/13.pdf>

34. Орлов А.И. Новая парадигма анализа статистических и экспертных данных в задачах экономики и управления / А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №04(098). С. 105 – 125. – IDA [article ID]: 0981404008. – Режим доступа: <http://ej.kubagro.ru/2014/04/pdf/08.pdf>

35. Орлов А.И. Системная нечеткая интервальная математика (СНИМ) – перспективное направление теоретической и вычислительной математики / А.И. Орлов, Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №07(091). С. 255 – 308. – IDA [article ID]: 0911307015. – Режим доступа: <http://ej.kubagro.ru/2013/07/pdf/15.pdf>

36. Орлов А.И., Луценко Е.В. Системная нечеткая интервальная математика. Монография (научное издание). – Краснодар, КубГАУ. 2014. – 600 с.

## References

1. Suppes P., Zines Dzh. Osnovy teorii izmerenij // Psihologicheskie izmereniya. - М.: Mir, 1967. - S.9 - 110.

2. Orlov A.I. Ustojchivost' v social'no-jekonomicheskikh modeljah. - М.: Nauka, 1979. - 296 s.

3. Beljaev Ju.K. Veroyatnostnye metody vyborochnogo kontrolja. - М.: Nauka, 1975. - 408 s.

4. Lumel'skij Ja.P. Statisticheskie ocenki rezul'tatov kontrolja kachestva. - М.: Izd-vo standartov, 1979. - 200 s.

5. Orlov A.I. Organizacionno-jekonomicheskoe modelirovanie : uchebnik : v 3 ch. Ch.3. Statisticheskie metody analiza dannyh. - М.: Izd-vo MGTU im. N.Je. Baumana, 2012. - 624 s.

6. Djevid G. Metod parnyh sravnenij. - M.: Statistika, 1978.- 144 s.
7. Orlov A.I. Organizacionno-jekonomicheskoe modelirovanie: ucheb. Ch.2. Jekspertnye ocenki. - M.: Izd-vo MGTU im. N.Je. Baumana, 2011. - 486 s.
8. Verojatnost' i matematicheskaja statistika: Jenciklopedija / Gl. red. Ju.V. Prohorov. - M.: Bol'shaja Rossijskaja jenciklopedija, 1999. - 910 s.
9. 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.
10. Orlov A.I. Parnye sravnenija v asimptotike Kolmogorova. – V sb.: Jekspertnye ocenki v zadachah upravlenija. - M.: Izd-vo Instituta problem upravlenija AN SSSR, 1982. - S. 58-66.
11. 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>
12. Orlov A.I. Statistika ob#ektov nechislovoj prirody i jekspertnye ocenki. – V sb.: Jekspertnye ocenki / Voprosy kibernetiki. Vyp.58. - M.: Nauchnyj Sovet AN SSSR po kompleksnoj probleme "Kibernetika", 1979. - S.17-33.
13. Orlov A.I. Sluchajnye mnozhestva s nezavisimymi jelementami (ljusiany) i ih primenenija. – V sb.: Algoritmicheskoe i programmnoe obespechenie prikladnogo statisticheskogo analiza. Uchenye zapiski po statistike, t.36. - M.: Nauka, 1980. - S. 287-308.
14. Rydanova G.V. Nekotorye voprosy statisticheskogo analiza sluchajnyh binarnyh vektorov. Diss. ... kand. fiz.-mat. nauk. - M.: MGU, f-t vychislitel'noj matematiki i kibernetiki, 1987. - 139 s.
15. Aksenova G.A., Kuz'mina E.S., Orlov A.I., Rozova N.K. Kinetotopografija v diagnostike infarkta miokarda. – V sb.: Aktual'nye voprosy klinicheskoy i jekspertnoj mediciny. – M.: 4 Glavnoe Upravlenie pri Minzdrave SSSR, 1979. S.24-26.
16. 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. - Zhurnal «Klinicheskaja medicina». 1982. T.LX. No.3. S.25-30.
17. Metodicheskie rekomendacii po provedeniju jekspertnoj ocenki planiruemyh i zakonchennyh nauchnyh rabot v oblasti mediciny (po problemam sojuznogo znachenija) / Sostaviteli: G.V. Raushenbah, O.V. Filippov. - M.: AMN SSSR - Uchenyj medicinskij sovet Minzdrava SSSR, 1982. - 36 s.
18. Leman Je. Proverka statisticheskikh gipotez. - M.: Nauka, 1979. - 408 s.
19. Borovkov A.A. Matematicheskaja statistika / Uchebnoe posobie dlja vuzov. - M.: Nauka, 1984. - 472 s.
20. Kendjel M. Rangovyje korreljacii. - M.: Statistika, 1975. - 216 s.
21. Bol'shev L.N., Smirnov N.V. Tablicy matematicheskoy statistiki. - M.: Nauka, 1965 (1-e izd.), 1968 (2-e izd.), 1983 (3-e izd.).
22. Orlov A.I. Problema mnozhestvennyh proverok statisticheskikh gipotez // Zavodskaja laboratorija. Diagnostika materialov. 1996. T.62. № 5. S.51 - 54.
23. Ljubishhev A.A. Dispersionnyj analiz v biologii. - M.: Izd-vo MGU, 1986. - 200 s.
24. Dyl'ko T.N. Proverka gipotez v jekspertnom ocenivanii // Vestnik Belorusskogo gosudarstvennogo universiteta / Ser. 1: Fizika, matematika i mehanika. 1988. № 2. S.36 - 40.
25. Orlov A.I., Raushenbah G.V. Metrika podobija: aksiomaticheskoe vvedenie, asimptoticheskaja normal'nost'. - V sb.: Statisticheskie metody ocenivanja i proverki gipotez. Mezhvuzovskij sbornik nauchnyh trudov. - Perm': Izd-vo Permskogo gosudarstvennogo

universiteta, 1986, s.148-157.

26. Orlov A.I. Jekonometrika. Uchebnik dlja vuzov. Izd. 3-e, ispravlennoe i dopolnennoe. - M.: Izd-vo «Jekzamen», 2004. – 576 s.

27. Orlov A.I., Mironova N.G., Fomin V.N., Chercincev A.N. Rekomendacii. Prikladnaja statistika. Metody obrabotki dannyh. Osnovnye trebovanija i harakteristiki. - M.: VNII standartizacii Gosstandarta SSSR, 1987. - 62 s.

28. Orlov A.I. Srednie velichiny i zakony bol'shih chisel v prostranstvah proizvol'noj prirody / A.I. Orlov // Politematicheskij setевой 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>

29. Orlov A.I. Asimptotika reshenij jekstremal'nyh statisticheskikh zadach // Analiz nechislovyh dannyh v sistemnyh issledovanijah. Sbornik trudov. Vyp.10. - M.: Vsesojuznyj nauchno-issledovatel'skij institut sistemnyh issledovanij, 1982. S.4 - 12.

30. Tjurin Ju.N., Vasilevich A.P. K probleme obrabotki rjadov ranzhirovok // Statisticheskie metody analiza jekspertnyh ocenok. Uchenye zapiski po statistike, t.29. - M.: Nauka, 1977. - S.96 - 111.

31. Orlov A.I. Nekotorye verojatnostnye voprosy teorii klassifikacii // Prikladnaja statistika. Uchenye zapiski po statistike, t.45. - M.: Nauka, 1983. - S.166-179.

32. Orlov A.I. O razvitii statistiki ob#ektov nechislovoj prirody / A.I. Orlov // Politematicheskij setевой 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>

33. 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. 187 – 213. – IDA [article ID]: 0901306013. – Rezhim dostupa: <http://ej.kubagro.ru/2013/06/pdf/13.pdf>

34. Orlov A.I. Novaja paradigma analiza statisticheskikh i jekspertnyh dannyh v zadachah jekonomiki i upravlenija / A.I. Orlov // Politematicheskij setевой jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2014. – №04(098). S. 105 – 125. – IDA [article ID]: 0981404008. – Rezhim dostupa: <http://ej.kubagro.ru/2014/04/pdf/08.pdf>

35. Orlov A.I. Sistemnaja nechetskaja interval'naja matematika (SNIM) – perspektivnoe napravlenie teoreticheskoy i vychislitel'noj matematiki / A.I. Orlov, E.V. Lucenko // Politematicheskij setевой jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2013. – №07(091). S. 255 – 308. – IDA [article ID]: 0911307015. – Rezhim dostupa: <http://ej.kubagro.ru/2013/07/pdf/15.pdf>

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