2.4. Описание неопределенностей с помощью теории нечеткости

2.4.1. Нечеткие множества

Пусть A - некоторое множество. Подмножество B множества A характеризуется своей характеристической функцией

$$m_B(x) = \begin{cases} 1, & x \in B, \\ 0, & x \notin B. \end{cases} \qquad (1)$$

Что такое нечеткое множество? Обычно говорят, что нечеткое подмножество C множества A характеризуется своей функцией принадлежности $m_C : A \to [0,1]$. Значение функции принадлежности в точке x показывает степень принадлежности этой точки нечеткому множеству. Нечеткое множество описывает неопределенность, соответствующую точке x – она одновременно и входит, и не входит в нечеткое множество C. За вхождение - $m_c(x)$ шансов, за второе – (1-$m_c(x)$) шансов.

Если функция принадлежности $m_C(x)$ имеет вид (1) при некотором B, то C есть обычное (четкое) подмножество A. Таким образом, теория нечетких множество является не менее общей математической дисциплиной, чем обычная теория множеств, поскольку обычные множества – частный случай нечетких. Соответственно можно ожидать, что теория нечеткости как целое обобщает классическую математику. Однако позже мы увидим, что теория нечеткости в определенном смысле сводится к теории случайных множеств и тем самым является частью классической математики. Другими словами, по степени общности обычная математика и нечеткая математика эквивалентны. Однако для практического применения в теории принятия решений описание и анализ неопределенностей с помощью теории нечетких множеств весьма плодотворны.

Обычное подмножество можно было бы отождествить с его характеристической функцией. Этого математики не делают, поскольку для задания функции (в ныне принятом подходе) необходимо сначала задать множество. Нечеткое же подмножество с формальной точки зрения можно отождествить с его функцией принадлежности. Однако термин "нечеткое подмножество" предпочтительнее при построении математических моделей реальных явлений.

Теория нечеткости является обобщением интервальной математики. Действительно, функция принадлежности

$$m_B(x) = \begin{cases} 1, & x \in [a,b], \\ 0, & x \notin [a,b] \end{cases}$$

задает интервальную неопределенность – про рассматриваемую величину известно лишь, что она лежит в заданном интервале $[a,b]$. Тем самым описание неопределенностей с помощью нечетких множеств является более общим, чем с помощью интервалов.

Начало современной теории нечеткости положено работой 1965 г. американского ученого азербайджанского происхождения Л.А.Заде. К настоящему времени по этой теории опубликованы тысячи книг и статей, издается несколько международных журналов, выполнено достаточно много как теоретических, так и прикладных работ. Первая книга российского автора по теории нечеткости вышла в 1980 г. [1].

Л.А. Заде рассматривал теорию нечетких множеств как аппарат анализа и моделирования гуманистических систем, т.е. систем, в которых участвует человек. Его подход опирается на предпосылку о том, что элементами мышления человека являются не числа, а элементы некоторых нечетких множеств или классов объектов, для которых переход от "принадлежности" к "непринадлежности" не скачкообразен, а непрерывен. В настоящее время методы теории нечеткости используются почти во всех прикладных областях, в том числе при управлении предприятием, качеством продукции и технологическими процессами.

Л.А. Заде использовал термин "fuzzy set" (нечеткое множество). На русский язык термин "fuzzy" переводили как нечеткий, размытый, расплывчатый, и даже как пушистый и туманный.

Аппарат теории нечеткости громоздок. В качестве примера дадим определения теоретико-множественных операций над нечеткими множествами. Пусть C и D- два нечетких подмножества A с функциями принадлежности $m_C(x)$ и $m_D(x)$ соответственно. Пересечением $C \cap D$, произведением CD, объединением $C \cup D$, отрицанием $\overline{C}$, суммой $C+D$ называются нечеткие подмножества A с функциями принадлежности

$$m_{C \cap D}(x) = \min(m_C(x), m_D(x)),\quad m_{CD}(x) = m_C(x) m_D(x),\quad m_{\overline{C}}(x) = 1 - m_C(x),$$

$$m_{C \cup D}(x) = \max(m_C(x), m_D(x)),\quad m_{C+D}(x) = m_C(x) + m_D(x) - m_C(x) m_D(x),\quad x \in A,$$

соответственно.

Как уже отмечалось, теория нечетких множеств в определенном смысле сводится к теории вероятностей, а именно, к теории случайных множеств. Соответствующий цикл теорем приведен ниже. Однако при решении прикладных задач вероятностно-статистические методы и методы теории нечеткости обычно рассматриваются как различные.

Для знакомства со спецификой нечетких множеств рассмотрим некоторые их свойства.

В дальнейшем считаем, что все рассматриваемые нечеткие множества являются подмножествами одного и того же множества Y.

Законы де Моргана для нечетких множеств. Как известно, законами же Моргана называются следующие тождества алгебры множеств

$$\overline{A\cup B}=\overline{A}\cap\overline{B},\quad \overline{A\cap B}=\overline{A}\cup\overline{B}. \qquad (2)$$

Теорема 1. Для нечетких множеств справедливы тождества

$$\overline{A\cup B}=\overline{A}\cap\overline{B},\quad \overline{A\cap B}=\overline{A}\cup\overline{B}, \qquad (3)$$

$$\overline{A+B}=\overline{A}\,\overline{B},\quad \overline{AB}=\overline{A}+\overline{B}. \qquad (4)$$

Доказательство теоремы 1 состоит в непосредственной проверке справедливости соотношений (3) и (4) путем вычисления значений функций принадлежности участвующих в этих соотношениях нечетких множеств на основе определений, данных выше.

Тождества (3) и (4) назовем *законами де Моргана для нечетких множеств*. В отличие от классического случая соотношений (2), они состоят из четырех тождеств, одна пара которых относится к операциям объединения и пересечения, а вторая - к операциям произведения и суммы. Как и соотношение (2) в алгебре множеств, законы де Моргана в алгебре нечетких множеств позволяют преобразовывать выражения и формулы, в состав которых входят операции отрицания.

Дистрибутивный закон для нечетких множеств. Некоторые свойства операций над множествами не выполнены для нечетких множеств. Так, $A+A\neq A$, за исключением случая, когда A - "четкое" множество (т.е. функция принадлежности принимает только значения 0 и 1).

Верен ли дистрибутивный закон для нечетких множеств? В литературе иногда расплывчато утверждается, что "не всегда". Внесем полную ясность.

Теорема 2. Для любых нечетких множеств A, B и C

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C). \qquad (5)$$

В то же время равенство

$$A(B+C) = AB + AC \qquad (6)$$

справедливо тогда и только тогда, когда при всех $y \in Y$

$$(m_A^2(y) - m_A(y))m_B(y)m_C(y) = 0.$$

Доказательство. Фиксируем произвольный элемент $y \in Y$. Для сокращения записи обозначим $a = m_A(y), b = m_B(y), c = m_C(y).$ Для доказательства тождества (5) необходимо показать, что

$$\min(a, \max(b, c)) = \max(\min(a, b), \min(a, c)). \qquad (7)$$

Рассмотрим различные упорядочения трех чисел a, b, c. Пусть сначала $a \le b \le c.$ Тогда левая часть соотношения (7) есть $\min(a, c) = a,$ а правая $\max(a, a) = a,$ т.е. равенство (7) справедливо.

Пусть $b \le a \le c.$ Тогда в соотношении (7) слева стоит $\min(a, c) = a,$ а справа $\max(b, a) = a,$ т.е. соотношение (7) опять является равенством.

Если $b \le c \le a,$ то в соотношении (7) слева стоит $\min(a, c) = c,$ а справа $\max(b, c) = c,$ т.е. обе части снова совпадают.

Три остальные упорядочения чисел a, b, c разбирать нет необходимости, поскольку в соотношение (6) числа b и c входят симметрично. Тождество (5) доказано.

Второе утверждение теоремы 2 вытекает из того, что в соответствии с определениями операций над нечеткими множествами

$$m_{A(B+C)}(y) = a(b + c - bc) = ab + ac - abc$$

и

$$m_{AB+AC}(y) = ab + ac - (ab)(ac) = ab + ac - a^2bc.$$

Эти два выражения совпадают тогда и только тогда, когда, когда $a^2bc = abc,$ что и требовалось доказать.

Определение 1. Носителем нечеткого множества *A* называется совокупность всех точек $y \in Y$, для которых $m_A(y) > 0$.

Следствие теоремы 2. Если носители нечетких множеств *B* и *C* совпадают с *У*, то равенство *(6) имеет место тогда и только тогда, когда A* - "четкое" (т.е. обычное, классическое, не нечеткое) множество.

Доказательство. По условию $m_B(y) m_C(y) \neq 0$ при всех $y \in Y$. Тогда из теоремы 2 следует, что $m_A^2(y) - m_A(y) = 0$, т.е. $m_A(y) = 1$ или $m_A(y) = 0$, что и означает, что *A* - четкое множество.

2.4.2. Пример описания неопределенности с помощью нечеткого множества

Понятие «богатый» часто используется при обсуждении социально-экономических проблем, в том числе и в связи с подготовкой и принятием решений. Однако очевидно, что разные лица вкладывают в это понятие различное содержание. Сотрудники Института высоких статистических технологий и эконометрики провели в 1996 г. социологическое исследование представления различных слоёв населения о понятии "богатый человек".

Мини-анкета опроса выглядела так:

1. При каком месячном доходе (в млн. руб. на одного человека) Вы считали бы себя богатым человеком?
2. Оценив свой сегодняшний доход, к какой из категорий Вы себя относите:

а) богатые;

б) достаток выше среднего;

в) достаток ниже среднего;

г) бедные;

д) за чертой бедности?

(В дальнейшем вместо полного наименования категорий будем оперировать буквами, например "в" - категория, "б" - категория и т.д.)

3. Ваша профессия, специальность.

Всего было опрошено 74 человека, из них 40 - научные работники и преподаватели, 34 человека - не занятых в сфере науки и образования, в том числе 5 рабочих и 5 пенсионеров. Из всех опрошенных только один (!) считает себя богатым. Несколько типичных ответов научных работников и преподавателей приведено в табл.1, а аналогичные сведения для работников коммерческой сферы – в табл.2.

Таблица 1.

Типичные ответы научных работников и преподавателей

Ответы на вопрос 3	Ответы на вопрос 1, млн. руб./чел.	Ответы на вопрос 2	Пол
Кандидат наук	1	д	ж
Преподаватель	1	в	ж
Доцент	1	б	ж
Учитель	10	в	м
Старший. научный сотрудник	10	д	м
Инженер-физик	24	д	ж
Программист	25	г	м
научный работник	45	г	м

Таблица 2

Типичные ответы работников коммерческой сферы.

Ответы на вопрос 3	Ответы на	Ответы на	Пол

	вопрос 1	вопрос 2	
Вице-президент банка	100	а	ж
Зам. директора банка	50	б	ж
Начальник. кредитного отдела	50	б	м
Начальник отдела ценных бумаг	10	б	м
Главный бухгалтер	20	д	ж
Бухгалтер	15	в	ж
Менеджер банка	11	б	м
Начальник отдела проектирования	10	в	ж

Разброс ответов на первый вопрос – от 1 до 100 млн. руб. в месяц на человека. Результаты опроса показывают, что критерий богатства у финансовых работников в целом несколько выше, чем у научных (см. гистограммы на рис.1 и рис.2 ниже).

Опрос показал, что выявить какое-нибудь конкретное значение суммы, которая необходима "для полного счастья", пусть даже с небольшим разбросом, нельзя, что вполне естественно. Как видно из таблиц 1 и 2, денежный эквивалент богатства колеблется от 1 до 100 миллионов рублей в месяц. Подтвердилось мнение, что работники сферы образования в подавляющем большинстве причисляют свой достаток к категории "в" и ниже (81% опрошенных), в том числе к категории "д" отнесли свой достаток 57%.

Со служащими коммерческих структур и бюджетных организаций иная картина: "г" - категория 1 человек (4%), "д" -

категория 4 человека (17%), "б" - категория - 46% и 1 человек "а" - категория.

Пенсионеры, что не вызывает удивления, отнесли свой доход к категории "д" (4 человека), и лишь один человек указал "г" - категорию. Рабочие же ответили так: 4 человека - "в", и один человек - "б".

Для представления общей картины в табл.3 приведены данные об ответах работников других профессий.

Таблица 3.

Типичные ответы работников различных профессий.

Ответы на вопрос 3	Ответы на вопрос 1	Ответы на вопрос 2	Пол
Работник торговли	1	б	ж
Дворник	2	в	ж
Водитель	10	в	м
Военнослужащий	10	в	м
Владелец бензоколонки	20	б	ж
Пенсионер	6	д	ж
Начальник фабрики	20	б	м
Хирург	5	в	м
Домохозяйка	10	в	ж
Слесарь-механик	25	в	м
Юрист	10	б	м
Оператор ЭВМ	20	д	м
Работник собеса	3	д	ж
Архитектор	25	б	ж

Прослеживается интересное явление: чем выше планка богатства для человека, тем к более низкой категории относительно этой планки он себя относит.

Для сводки данных естественно использовать гистограммы. Для этого необходимо сгруппировать ответы. Использовались 7 классов (интервалов):

1 – до 5 миллионов рублей в месяц на человека (включительно);

2 – от 5 до 10 миллионов;

3- от 10 до 15 миллионов;

4 – от 15 до 20 миллионов;

5 – от 20 до 25 миллионов;

6 – от 25 до 30 миллионов;

7 – более 30 миллионов.

(Во всех интервалах левая граница исключена, а правая, наоборот – включена.)

Сводная информация представлена на рис.1 (для научных работников и преподавателей) и рис.2 (для всех остальных, т.е. для лиц, не занятых в сфере науки и образования - служащих иных бюджетных организаций, коммерческих структур, рабочих, пенсионеров).

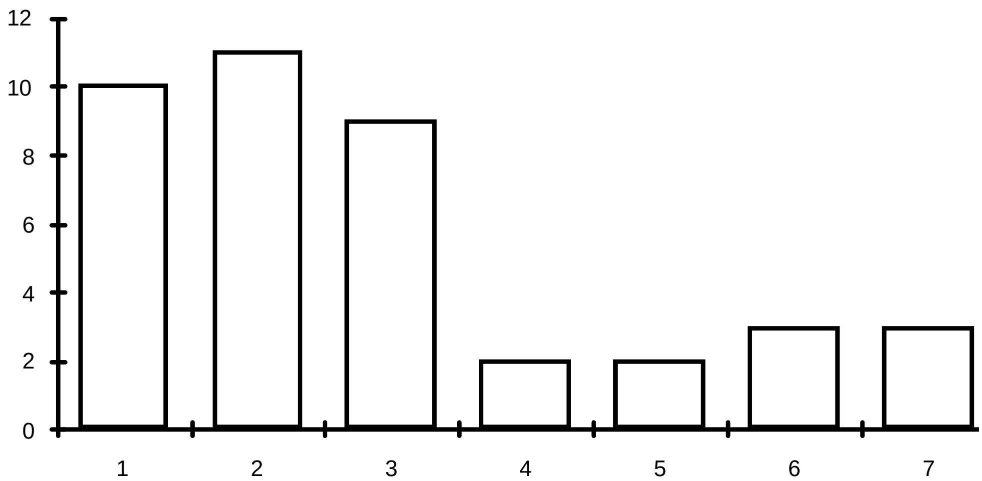

Рис.1. Гистограмма ответов на вопрос 1 для научных работников и преподавателей (40 человек).

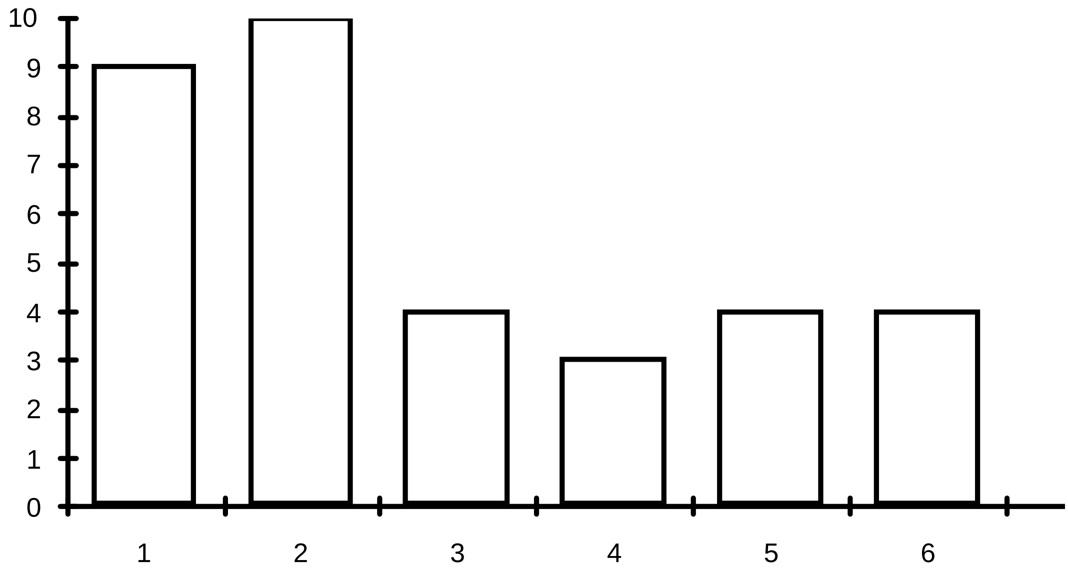

Рис.2. Гистограмма ответов на вопрос 1 для лиц, не занятых в сфере науки и образования (34 человека).

Для двух выделенных групп, а также для некоторых подгрупп второй группы рассчитаны сводные средние характеристики – выборочные средние арифметические, медианы, моды. При этом медиана группы - количество млн. руб., названное центральным по порядковому номеру опрашиваемым в возрастающем ряду ответов на вопрос 1, а мода группы - интервал, на котором столбик гистограммы - самый высокий, т.е. в него "попало" максимальное количество опрашиваемых. Результаты приведены в табл. 4.

Таблица 4.

Сводные средние характеристики ответов на вопрос 1 для различных групп (в млн. руб. в мес. на чел.).

Группа опрошенных	Среднее арифметическое	медиана	мода
Научные работники и преподаватели	11,66	7,25	(5; 10)
Лиц, не занятых в сфере науки и образования	14,4	20	(5; 10)
Служащие коммерческих структур и бюджетных организаций	17,91	10	(5; 10)
Рабочие	15	13	-
Пенсионеры	10,3	10	-

Построим нечеткое множество, описывающее понятие «богатый человек» в соответствии с представлениями опрошенных. Для этого составим табл.5 на основе рис.1 и рис.2 с учетом размаха ответов на первый вопрос.

Таблица 5.

Число ответов, попавших в интервалы

№	Номер интервала	0	1	2	3	4
1	Интервал, млн. руб. в месяц	(0;1)	[1;5]	(5;10]	(10;15]	(15;20]
2	Число ответов в интервале	0	19	21	13	5
3	Доля ответов в интервале	0	0,257	0,284	0,176	0,068
4	Накопленное число ответов	0	19	40	53	58
5	Накопленная доля ответов	0	0,257	0,541	0,716	0,784

Продолжение табл.5.

№	Номер интервала	5	6	7	8
1	Интервал, млн. руб. в месяц	(20;25]	(25;30]	(30;100)	[100;+∞)
2	Число ответов в интервале	6	7	2	1
3	Доля ответов в интервале	0,081	0,095	0,027	0,013
4	Накопленное число ответов	64	71	73	74
5	Накопленная доля ответов	0,865	0,960	0,987	1,000

Пятая строка табл.5 задает функцию принадлежности нечеткого множества, выражающего понятие "богатый человек" в терминах его ежемесячного дохода. Это нечеткое множество является подмножеством множества из 9 интервалов, заданных в строке 2 табл.5. Или множества из 9 условных номеров {0, 1, 2, …, 8}. Эмпирическая функция распределения, построенная по выборке из ответов 74 опрошенных на первый вопрос мини-анкеты, описывает понятие "богатый человек" как нечеткое подмножество положительной полуоси.

2.4.3. О разработке методики ценообразования на основе теории нечетких множеств

Для оценки значений показателей, не имеющих количественной оценки, можно использовать методы нечетких множеств. Например, в диссертации П.В. Битюкова [2] нечеткие множества применялись при моделировании задач ценообразования на электронные обучающие курсы, используемые при дистанционном обучении. Им было проведено исследование значений фактора «Уровень качества курса» с

использованием нечетких множеств. В ходе практического использования предложенной П.В. Битюковым методики ценообразования значения ряда других факторов могут также определяться с использованием теории нечетких множеств. Например, ее можно использовать для определения прогноза рейтинга специальности в вузе с помощью экспертов, а также значений других факторов, относящихся к группе «Особенности курса». Опишем подход П.В. Битюкова как пример практического использования теории нечетким множеств.

Значение оценки, присваиваемой каждому интервалу для фактора «Уровень качества курса», определяется на универсальной шкале [0,1], где необходимо разместить значения лингвистической переменной «Уровень качества курса»: НИЗКИЙ, СРЕДНИЙ, ВЫСОКИЙ. Степень принадлежности некоторого значения вычисляется как отношение числа ответов, в которых оно встречалось в определенном интервале шкалы, к максимальному (для этого значения) числу ответов по всем интервалам.

В ходе работы над диссертацией был проведен опрос экспертов о степени влияния уровня качества электронных курсов на их потребительную ценность. Каждому эксперту в процессе опроса предлагалось оценить с позиции потребителя ценность того или иного класса курсов в зависимости от уровня качества. Эксперты давали свою оценку для каждого класса курсов по 10-ти балльной шкале (где 1 - min, 10 - max). Для перехода к универсальной шкале [0,1], все значения 10-ти балльной шкалы оценки ценности были разделены на максимальную оценку 10.

Используя свойства функции принадлежности, необходимо предварительно обработать данные с тем, чтобы уменьшить искажения, вносимые опросом. Естественными свойствами функций принадлежности являются наличие одного максимума и гладкие, затухающие до нуля фронты. Для обработки статистических данных можно воспользоваться так называемой матрицей подсказок. Предварительно удаляются явно ошибочные элементы. Критерием удаления служит наличие нескольких нулей в строке вокруг этого элемента.

Элементы матрицы подсказок вычисляются по формуле:

$$k_j = \sum_{i=1}^{n} b_{ij},\ j = \overline{1,n},$$

где b_{ij} - элемент таблицы с результатами анкетирования, сгруппированными по интервалам. Матрица подсказок представляет собой строку, в которой выбирается максимальный элемент: $k_{\max} = \max_j k_j$, и далее все ее элементы преобразуются по формуле:

$$c_{ij} = \frac{b_{ij} k_{\max}}{k_j},\ i = \overline{1,m},\ j = \overline{1,n}.$$

Для столбцов, где $k_j = 0$, применяется линейная аппроксимация:

$$c_{ij} = \frac{c_{ij-1} + c_{ij+1}}{2},\ i = \overline{1,m},\ j = \overline{1,n}.$$

Результаты расчетов сводятся в таблицу, на основании которой строятся функции принадлежности. Для этого находятся максимальные элементы по строкам: $c_{i\max} = \max_j c_{ij},\ i = \overline{1,m},\ j = \overline{1,n}$. Функция принадлежности вычисляется по формуле: $\mu_{ij} = c_{ij} / c_{i\max}$. Результаты расчетов приведены в табл. 6.

Таблица 6

Значения функции принадлежности лингвистической переменной

μ_i	Интервал на универсальной шкале

	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
μ_1	0	0,2	1	1	0,89	0,67	0	0	0	0
μ_2	**0**	**0**	**0**	**0**	**0**	**0,33**	**1**	**1**	**0**	**0**
μ_3	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**1**

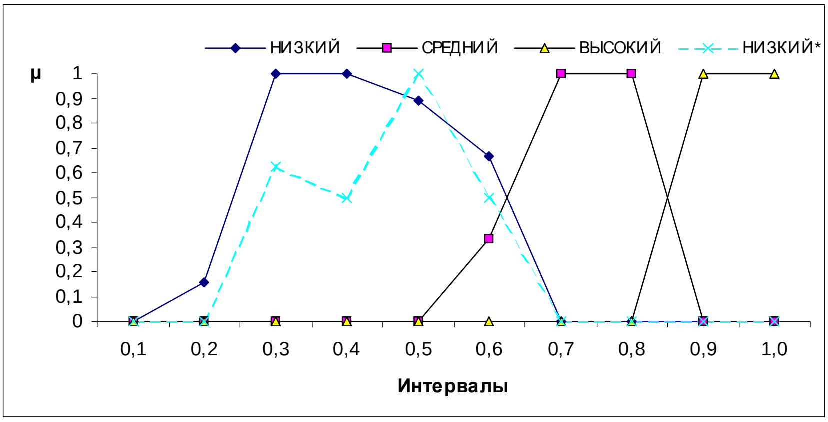

Рис. 3. График функций принадлежности значений лингвистической переменной «Уровень качества курса»

На рис.3 сплошными линиями показаны функции принадлежности значений лингвистической переменной «Уровень качества курса» после обработки таблицы, содержащей результаты опроса. Как видно из графика, функции принадлежности удовлетворяют описанным выше свойствам. Для сравнения пунктирной линией показана функция принадлежности лингвистической переменной для значения НИЗКИЙ без обработки данных.

2.4.4. О статистике нечетких множеств

Нечеткие множества – частный вид объектов нечисловой природы. Статистические методы анализа объектов нечисловой

природы описаны в [3]. В частности, среднее значение нечеткого множества можно определить по формуле:

$$M(A)=\frac{\sum_{i=1}^{n} x_i \mu_A(x_i)}{\sum_{i=1}^{n} \mu_A(x_i)},$$

где $\mu_A(x_i)$**- функция принадлежности нечеткого множества** ***A*****.**

Как известно, методы статистики нечисловых данных базируются на использовании расстояний (или показателей различия) в соответствующих пространствах нечисловой природы. Расстояние между нечеткими подмножествами ***A*** **и** ***B*** **множества** $X = \{x_1, x_2, \ldots, x_k\}$ **можно определить как**

$$d(A,B)=\sum_{j=1}^{k} | m_A(x_j) - m_B(x_j) |,$$

где $\mu_A(x_j)$**- функция принадлежности нечеткого множества** ***A*****, а** $\mu_B(x_i)$**- функция принадлежности нечеткого множества** ***B*****. Может использоваться и другое расстояние:**

$$d_1(A,B)=\frac{\sum_{j=1}^{k} | m_A(x_j) - m_B(x_j) |}{\sum_{j=1}^{k} (m_A(x_j) + m_B(x_j))}.$$

(Примем это расстояние равным 0, если функции принадлежности тождественно равны 0.)

В соответствии с аксиоматическим подходом к выбору расстояний (метрик) в пространствах нечисловой природы разработан обширный набор систем аксиом, из которых выводится тот или иной вид расстояний (метрик) в конкретных пространствах [1, 3, 4]. При использовании вероятностных моделей расстояние между случайными нечеткими множествами

само является случайной величиной, имеющей в ряде постановок асимптотически нормальное распределение [5].

2.4.5. Нечеткие множества как проекции случайных множеств

С самого начала появления современной теории нечеткости в 1960-е годы началось обсуждение ее взаимоотношений с теорией вероятностей. Дело в том, что функция принадлежности нечеткого множества напоминает распределение вероятностей. Отличие только в том, что сумма вероятностей по всем возможным значениям случайной величины (или интеграл, если множество возможных значений несчетно) всегда равна 1, а сумма S значений функции принадлежности (в непрерывном случае - интеграл от функции принадлежности) может быть любым неотрицательным числом. Возникает искушение пронормировать функцию принадлежности, т.е. разделить все ее значения на S (при $S \neq 0$), чтобы свести ее к распределению вероятностей (или к плотности вероятности). Однако специалисты по нечеткости справедливо возражают против такого "примитивного" сведения", поскольку оно проводится отдельно для каждой размытости (нечеткого множества), и определения обычных операций над нечеткими множествами с ним согласовать нельзя. Последнее утверждение означает следующее. Пусть указанным образом преобразованы функции принадлежности нечетких множеств A и B. Как при этом преобразуются функции принадлежности $A \cap B, A \cup B, A + B, AB$? Установить это *невозможно в принципе*. Последнее утверждение становится совершенно ясным после рассмотрения нескольких примеров пар нечетких множеств с одними и теми же суммами значений функций принадлежности, но различными результатами теоретико-множественных операций над ними, причем и суммы значений соответствующих функций принадлежности для этих результатов теоретико-множественных операций, например, для пересечений множеств, также различны.

В работах по нечетким множествам довольно часто утверждается, что теория нечеткости является самостоятельным разделом прикладной математики и не имеет отношения к теории вероятностей (см., например, обзор литературы в монографиях [1,4]). Авторы, сравнивавшие теорию нечеткости и теорию вероятностей, обычно подчеркивали различие между этими областями теоретических и прикладных исследований. Обычно сравнивают аксиоматику и сравнивают области приложений. Надо сразу отметить, что аргументы при втором типе сравнений не имеют доказательной силы, поскольку по поводу границ применимости даже такой давно выделившейся научной области, как вероятностно-статистические методы, имеются различные мнения. Напомним, что итог рассуждений одного из наиболее известных французских математиков Анри Лебега по поводу границ применимости арифметики таков: "Арифметика применима тогда, когда она применима" (см. его монографию [6, с.21-22]).

При сравнении различных аксиоматик теории нечеткости и теории вероятностей нетрудно увидеть, что списки аксиом различаются. Из этого, однако, отнюдь не следует, что между указанными теориями нельзя установить связь, типа известного сведения евклидовой геометрии на плоскости к арифметике (точнее к теории числовой системы R^2 - см., например, монографию [7]). Напомним, что эти две аксиоматики - евклидовой геометрии и арифметики - на первый взгляд весьма сильно различаются.

Можно понять желание энтузиастов нового направления подчеркнуть принципиальную новизну своего научного аппарата. Однако не менее важно установить связи нового подхода с ранее известными.

Как оказалось, теория нечетких множеств тесно связана с теорией случайных множеств. Еще в 1974 г. в работе [8] было показано, что нечеткие множества естественно рассматривать как "проекции" случайных множеств. Рассмотрим этот метод сведения теории нечетких множеств к теории случайных множеств.

Определение 2. Пусть $A = A(w)$ - *случайное подмножество конечного множества У*. Нечеткое множество *В*, определенное на *У*, называется проекцией *А* и обозначается *Proj A*, если

$$m_B(y) = P(y \in A) \qquad (8)$$

при всех $y \in Y$.

Очевидно, каждому случайному множеству *А* можно поставить в соответствие с помощью формулы (8) нечеткое множество *В* = *Proj A*. Оказывается, верно и обратное.

Теорема 3. Для любого нечеткого подмножества *В* конечного множества *У* существует случайное подмножество *А* множества *У* такое, что *В* = *Proj A*.

Доказательство. Достаточно задать распределение случайного множества *А*. Пусть $У_1$ - носитель *В* (см. определение 1 выше). Без ограничения общности можно считать, что $Y_1 = \{y_1, y_2, ..., y_m\}$ при некотором *m* и элементы $У_1$ занумерованы в таком порядке, что

$$0 < m_B(y_1) \le m_B(y_2) \le ... \le m_B(y_m).$$

Введем множества

$$Y(1) = Y_1, Y(2) = \{y_2, ..., y_m\}, ..., Y(t) = \{y_t, ..., y_m\}, ..., Y(m) = \{y_m\}.$$

Положим

$$P(A=Y(1))=m_B(y_1),\quad P(A=Y(2))=m_B(y_2)-m_B(y_1),...,$$

$$P(A=Y(t))=m_B(y_t)-m_B(y_{t-1}),...,P(A=Y(m))=m_B(y_m)-m_B(y_{m-1}),$$

$$P(A=\varnothing)=1-m_B(y_m).$$

Для всех остальных подмножеств X множества $У$ положим $P(A=X)=0$. Поскольку элемент y_t входит во множества $Y(1), Y(2),\dots, Y(t)$ и не входит во множества $Y(t+1),\dots, Y(m)$, то из приведенных выше формул следует, что $P(y_t\in A)=m_B(y_t)$. Если $y\notin Y_1$, то, очевидно, $P(y\in A)=0$. Теорема 3 доказана.

Распределение случайного множества с независимыми элементами, как следует из рассмотрений главы 8 монографии [3], полностью определяется его проекцией. Для конечного случайного множества общего вида это не так. Для уточнения сказанного понадобится следующая теорема.

Теорема 4. Для случайного подмножества A множества $У$ из конечного числа элементов наборы чисел $P(A=X), X\subseteq Y$, и $P(X\subseteq A), X\subseteq Y$, выражаются один через другой.

Доказательство. Второй набор выражается через первый следующим образом:

$$P(X\subseteq A)=\sum_{X':X\subseteq X'}P(A=X').$$

Элементы первого набора выразить через второй можно с помощью формулы включений и исключений из формальной логики, в соответствии с которой

$$P(A=X)=P(X\subseteq A)-\sum P(X\cup\{y\}\subseteq A)+\sum P(X\cup\{y_1,y_2\}\subseteq A)-...\pm P(Y\subseteq A).$$

В этой формуле в первой сумме y пробегает все элементы множества $Y\backslash X$, во второй сумме переменные суммирования y_1 и y_2 не совпадают и также пробегают это множество, и т.д. Ссылка на формулу включений и исключений завершает доказательство теоремы 4.

В соответствии с теоремой 4 случайное множество А можно характеризовать не только распределением, но и набором чисел $P(X \subseteq A), X \subseteq Y$. В этом наборе $P(\varnothing \subseteq A) = 1$, а других связей типа равенств нет. В этот набор входят числа $P(\{y\} \subseteq A) = P(y \in A)$, следовательно, фиксация проекции случайного множества эквивалентна фиксации $k = Card(Y)$ параметров из $(2^k\text{-}1)$ параметров, задающих распределение случайного множества A в общем случае.

Будет полезна следующая теорема.

Теорема 5. *Если* $Proj\ A = B$, *то* $\Pr oj\overline{A} = \overline{B}$.

Для доказательства достаточно воспользоваться тождеством из теории случайных множеств $P(\overline{A} = X) = P(A = \overline{X})$, формулой для вероятности накрытия $P(y \in A)$, определением отрицания нечеткого множества и тем, что сумма всех $P(A=X)$ равна 1. При этом под формулой для вероятности накрытия имеется в виду следующее утверждение: чтобы найти вероятность накрытия фиксированного элемента q случайным подмножеством S конечного множества Q, достаточно вычислить

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

где суммирование идет по всем подмножествам A множества Q, содержащим q.

2.4.6. Пересечения и произведения нечетких и случайных множеств

Выясним, как операции над случайными множествами соотносятся с операциями над их проекциями. В силу законов де Моргана (теорема 1) и теоремы 5 достаточно рассмотреть операцию пересечения случайных множеств.

Теорема 6. Если случайные подмножества A_1 и A_2 конечного множества $У$ независимы, то нечеткое множество $\Pr oj(A_1 \cap A_2)$ является произведением нечетких множеств *Proj* A_1 и *Proj* A_2.

Доказательство. Надо показать, что для любого $y \in Y$

$$P(y \in A_1 \cap A_2) = P(y \in A_1)P(y \in A_2). \qquad (9)$$

По формуле для вероятности накрытия точки случайным множеством (см. выше)

$$P(y \in A_1 \cap A_2) = \sum_{X : y \in X} P((A_1 \cap A_2) = X). \qquad (10)$$

Легко проверить, что распределение пересечения случайных множеств $A_1 \cap A_2$ можно выразить через их совместное распределение следующим образом:

$$P(A_1 \cap A_2 = X) = \sum_{X_1, X_2 : X_1 \cap X_2 = X} P(A_1 = X_1, A_2 = X_2). \qquad (11)$$

Из соотношений (10) и (11) следует, что вероятность накрытия для пересечения случайных множеств можно представить в виде двойной суммы

$$P(y \in A_1 \cap A_2) = \sum_{X : y \in X} \sum_{X_1, X_2 : X_1 \cap X_2 = X} P(A_1 = X_1, A_2 = X_2). \qquad (12)$$

Заметим теперь, что правую часть формулы (12) можно переписать следующим образом:

$$\sum_{X_1, X_2 : e \in X_1, e \in X_2} P(A_1 = X_1, A_2 = X_2). \qquad (13)$$

Действительно, формула (12) отличается от формулы (13) лишь тем, что в ней сгруппированы члены, в которых пересечение переменных суммирования $X_1 \cap X_2$ принимает постоянное значение. Воспользовавшись определением независимости случайных множеств и правилом перемножения сумм, получаем, что из (12) и (13) вытекает равенство

$$P(y\in A_1\cap A_2)=\left(\sum_{X_1:y\in X_1}P(A_1=X_1)\right)\left(\sum_{X_2:y\in X_2}P(A_2=X_2)\right).$$

Для завершения доказательства теоремы 6 достаточно еще раз сослаться на формулу для вероятности накрытия точки случайным множеством.

Определение 3. Носителем случайного множества C называется совокупность всех тех элементов $y\in Y,$ для которых $P(y\in C)>0.$

Теорема 7. Равенство

$$\mathrm{Pr}\,oj(A_1\cap A_2)=(\mathrm{Pr}\,ojA_1)\cap(\mathrm{Pr}\,ojA_2)$$

верно тогда и только тогда, когда пересечение носителей случайных множеств $\overline{A_1}\cap A_2$ и $A_1\cap\overline{A_2}$ пусто.

Доказательство. Необходимо выяснить условия, при которых

$$P(y\in A_1\cap A_2)=\min(P(y\in A_1),P(y\in A_2)). \qquad (14)$$

Положим

$$p_1=P(y\in A_1\cap A_2),\ p_2=P(y\in\overline{A_1}\cap A_2),\ p_3=P(y\in A_1\cap\overline{A_2}).$$

Тогда равенство (14) сводится к условию

$$p_1=\min(p_1+p_2,p_1+p_3). \quad (15)$$

Ясно, что соотношение (15) выполнено тогда и только тогда, когда $p_2p_3=0$ при всех $y\in Y,$ т.е. не существует ни одного элемента $y_0\in Y$ такого, что одновременно $P(y_0\in\overline{A_1}\cap A_2)>0$ и $P(y_0\in A_1\cap\overline{A_2})>0$, а это эквивалентно пустоте пересечения носителей случайных множеств $\overline{A_1}\cap A_2$ и $A_1\cap\overline{A_2}$. Теорема 7 доказана.

24.7. Сведение последовательности операций над нечеткими множествами к последовательности операций над случайными множествами

Выше получены некоторые связи между нечеткими и случайными множествами. Стоит отметить, что изучение этих связей в работе [8] началось с введения случайных множеств с целью развития и обобщения аппарата нечетких множеств Л. Заде. (Для фиксации приоритета на мировом уровне целесообразно отметить, что эта работа выполнена в 1974 г. и доложена в Центральном экономико-математическом институте АН СССР на всесоюзном научном семинаре "Многомерный статистический анализ и вероятностное моделирование реальных процессов" 18 декабря 1974 г. - см. [8, с.169].) Дело в том, что математический аппарат нечетких множеств не позволяет в должной мере учитывать различные варианты зависимости между понятиями (объектами), моделируемыми с его помощью, не является достаточно гибким. Так, для описания "общей части" двух нечетких множеств есть лишь две операции - произведение и пересечение. Если применяется первая из них, то фактически предполагается, что множества ведут себя как проекции независимых случайных множеств (см. выше теорему 6). Операция пересечения также накладывает вполне определенные ограничения на вид зависимости между множествами (см. выше теорему 7), причем в этом случае найдены даже необходимые и достаточные условия. Желательно иметь более широкие возможности для моделирования зависимости между множествами (понятиями, объектами). Использование математического аппарата случайных множеств предоставляет такие возможности.

Цель сведения теории нечетких множеств к теории случайных множеств состоит в том, чтобы за любой конструкцией из нечетких множеств увидеть конструкцию из случайных множеств, определяющую свойства первой, аналогично тому, как за плотностью распределения вероятностей мы видим случайную величину. В настоящем пункте приводим результаты по сведению алгебры нечетких множеств к алгебре случайных множеств.

Определение 4. Вероятностное пространство {Ω, *G, P*} назовем делимым, если для любого измеримого множества $X \in G$ и любого положительного числа a, меньшего $P(X)$, можно указать измеримое множество $Y \subset X$ такое, что $P(Y) = a$.

Пример. Пусть Ω - единичный куб конечномерного линейного пространства, *G* есть сигма-алгебра борелевских множеств, а *P* - мера Лебега. Тогда {Ω, *G, P*} - делимое вероятностное пространство.

Таким образом, делимое вероятностное пространство - это не экзотика. Обычный куб является примером такого пространства.

Доказательство сформулированного в примере утверждения проводится стандартными математическими приемами. Они основаны на том, что измеримое множество можно сколь угодно точно приблизить открытыми множествами, последние представляются в виде суммы не более чем счетного числа открытых шаров, а для шаров делимость проверяется непосредственно (от шара X тело объема $a < P(X)$ отделяется соответствующей плоскостью).

Теорема 8. Пусть даны случайное множество *А* на делимом вероятностном пространстве {Ω, *G, P*} со значениями во множестве всех подмножеств множества *У* из конечного числа элементов, и нечеткое множество *D* на *У*. Тогда существуют случайные множества C_1, C_2, C_3, C_4 на том же вероятностном пространстве такие, что

$$\mathrm{Pr}oj(A \cap C_1) = B \cap D, \quad \mathrm{Pr}oj(A \cap C_2) = BD, \quad \mathrm{Pr}oj(A \cup C_3) = B \cup D,$$

$$\mathrm{Pr}oj(A \cup C_4) = B + D, \quad \mathrm{Pr}ojC_i = D, \; i = 1,2,3,4,$$

где $B = Proj\, A$.

Доказательство. В силу справедливости законов де Моргана для нечетких (см. теорему 1 выше) и для случайных множеств, а также теоремы 5 выше (об отрицаниях) достаточно доказать существование случайных множеств C_1 и C_2.

Рассмотрим распределение вероятностей во множестве всех подмножеств множества Y, соответствующее случайному множеству C такому, что $Proj\ C = D$ (оно существует в силу теоремы 3). Построим случайное множество C_2 с указанным распределением, независимое от A. Тогда $\mathrm{Pr}oj(A \cap C_2) = BD$ по теореме 6.

Перейдем к построению случайного множества C_1. По теореме 7 необходимо и достаточно определить случайное множество $C_1(w)$ так, чтобы $ProjC_1 = D$ и пересечение носителей случайных множеств $A \cap \overline{C_1}$ и $\overline{A} \cap C_1$ было пусто, т.е.

$$p_3 = P(y \in A \cap \overline{C_1}) = 0$$

для $y \in Y_1 = \{y : m_B(y) \leq m_D(y)\}$ и

$$p_2 = P(y \in \overline{A} \cap C_1) = 0$$

для $y \in Y_2 = \{y : m_B(y) \geq m_D(y)\}$.

Построим $C_1(w)$, исходя из заданного случайного множества $A(w)$. Пусть $y_1 \in Y_2$. Исключим элемент y_1 из $A(w)$ для стольких элементарных событий w, чтобы для полученного случайного множества $A_1(w)$ было справедливо равенство

$$P(y_1 \in A_1) = m_D(y_1)$$

(именно здесь используется делимость вероятностного пространства, на котором задано случайное множество $A(w)$). Для $y \neq y_1$, очевидно,

$$P(y \in A_1) = P(y \in A).$$

Аналогичным образом последовательно исключаем y из $A(w)$ для всех $y \in Y_2$ и добавляем y в $A(w)$ для всех $y \in Y_1$, меняя на каждом шагу $P(y \in A_i)$ только для $y = y_i$ так, чтобы

$$P(y_i \in A_i) = m_D(y_i)$$

(ясно, что при рассмотрении $y_i \in Y_1 \cap Y_2$ случайное множество $A_i(w)$ не меняется). Перебрав все элементы $У$, получим случайное множество $A_k(w) = C_1(w)$, для которого выполнено требуемое. Теорема 8 доказана.

Основной результат о сведении теории нечетких множеств к теории случайных множеств дается следующей теоремой.

Теорема 9. Пусть $B_1, B_2, B_3, \ldots, B_t$ - некоторые нечеткие подмножества множества $У$ из конечного числа элементов. Рассмотрим результаты последовательного выполнения теоретико-множественных операций

$$B^m = ((\ldots((B_1 \circ B_2) \circ B_3) \circ \ldots) \circ B_{m-1}) \circ B_m, \quad m = 1,2,\ldots,t,$$

где $\circ$ - символ одной из следующих теоретико-множественных операций над нечеткими множествами: пересечение, произведение, объединение, сумма (на разных местах могут стоять разные символы). Тогда существуют случайные подмножества $A_1, A_2, A_3, \ldots, A_t$ того же множества $У$ такие, что

$$\mathrm{Pr}\,ojA_i = B_i, \; i = 1,2,\ldots,t,$$

и, кроме того, результаты теоретико-множественных операций связаны аналогичными соотношениями

$$\mathrm{Pr}\,oj\{((\ldots((A_1 \otimes A_2) \otimes A_3) \otimes \ldots) \otimes A_{m-1}) \otimes A_m\} = B^m, \; m = 1,2,\ldots,t,$$

где знак $\otimes$ означает, что на рассматриваемом месте стоит символ пересечения $\cap$ случайных множеств, если в определении B^m стоит символ пересечения или символ произведения нечетких множеств, и соответственно символ объединения $\cup$ случайных множеств, если в B^m стоит символ объединения или символ суммы нечетких множеств.

Комментарий. Поясним содержание теоремы. Например, если

$$B^5 = (((B_1 + B_2) \cap B_3) B_4) \cup B_5,$$

то

$$(((A_1 \otimes A_2) \otimes A_3) \otimes A_4) \otimes A_5 = (((A_1 \cup A_2) \cap A_3) \cap A_4) \cup A_5 .$$

Как совместить справедливость дистрибутивного закона для случайных множеств (вытекающего из его справедливости для обычных множеств) с теоремой 2 выше, в которой показано, что для нечетких множеств, вообще говоря, $(B_1 + B_2)B_3 \neq B_1B_3 + B_2B_3$? Дело в том, что хотя в соответствии с теоремой 9 для любых трех нечетких множеств B_1, B_2 и B_3 можно указать три случайных множества A_1, A_2 и A_3 такие, что

$$\mathrm{Pr}oj(A_i) = B_i,\ \ i = 1,2,3,\ \ \mathrm{Pr}oj(A_1 \cup A_2) = B_1 + B_2,\ \ \mathrm{Pr}oj((A_1 \cup A_2) \cap A_3) = B^3,$$

где

$$B^3 = (B_1 + B_2)B_3,$$

но при этом, вообще говоря,

$$\mathrm{Pr}oj(A_1 \cap A_3) \neq B_1B_3$$

и, кроме случаев, указанных в теореме 2,

$$\mathrm{Pr}oj((A_1 \cup A_2) \cap A_3) \neq B_1B_3 + B_2B_3 .$$

Доказательство теоремы 9 проводится по индукции. При t=1 распределение случайного множества строится с помощью теоремы 3. Затем конструируется само случайное множество A_1, определенное на делимом вероятностном пространстве (нетрудно проверить, что на делимом вероятностном пространстве можно построить случайное подмножество конечного множества с любым заданным распределением именно в силу делимости пространства). Далее случайные множества A_2, A_3, …, A_t строим по индукции с помощью теоремы 8. Теорема 9 доказана.

Замечание. Проведенное доказательство теоремы 9 проходит и в случае, когда при определении B^m используются отрицания, точнее, кроме B^m ранее введенного вида используются также последовательности результатов теоретико-множественных операций, очередной шаг в которых имеет вид

$$B_1^m = \overline{B^{m-1}} \circ B_m, \quad B_2^m = B^{m-1} \circ \overline{B_m}, \quad B_3^m = \overline{B^{m-1}} \circ \overline{B_m}.$$

А именно, сначала при помощи законов де Моргана (теорема 1 выше) проводится преобразование, в результате которого в последовательности B^m остаются только отрицания отдельных подмножеств из совокупности $B_1, B_2, B_3,..., B_t$, а затем с помощью теоремы 5 вообще удается избавиться от отрицаний и вернуться к условиям теоремы 9.

Итак, в настоящей главе описаны связи между такими объектами нечисловой природы, как нечеткие и случайные множества, установленные в нашей стране в первой половине 1970-х годов. Через несколько лет, а именно, в начале 1980-х годов, близкие подходы стали развиваться и за рубежом. Одна из работ [9] носит примечательное название "Нечеткие множества как классы эквивалентности случайных множеств".

В эконометрике [3] разработан ряд методов статистического анализа нечетких данных, в том числе методы классификации, регрессии, проверки гипотез о совпадении функций принадлежности по опытным данным и т.д., при этом оказались полезными общие подходы статистики объектов нечисловой природы (см. главу 8 в [3] и работы [1,4,8]). Методологические и прикладные вопросы теории нечеткости обсуждаются в литературе, в частности, в работах [1,4,10].

Литература

1. Орлов А.И. Задачи оптимизации и нечеткие переменные. – М.: Знание, 1980. - 64 с.

2. Битюков П.В. Моделирование задач ценообразования на электронные обучающие курсы в области дистанционного обучения / Автореферат диссертации на соискание ученой степени кандидата экономических наук. – М.: Московский государственный университет экономики, статистики и информатики, 2002. – 24 с.

3. Орлов А.И. Эконометрика. – М.: Экзамен, 2002. – 576 с.

4. Орлов А.И. Устойчивость в социально-экономических моделях. – М.: Наука, 1979. -296 с.

5. 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.

6. Лебег А. Об измерении величин. - М.: Учпедгиз, 1960. - 204 с.

7. Ефимов Н.В. Высшая геометрия. - М.: ГИФМЛ, 1961. - 580 с.

8. Орлов А.И. Основания теории нечетких множеств (обобщение аппарата Заде). Случайные толерантности. – В сб.: Алгоритмы многомерного статистического анализа и их применения. - М.: Изд-во ЦЭМИ АН СССР, 1975. - С.169-175.

9. Goodman I.R. Fuzzy sets as equivalence classes of random sets // Fuzzy Set and Possibility Theory: Recent Developments. - New York-Oxford-Toronto-Sydney-Paris-Frankfurt, Pergamon Press, 1982. - P.327-343. (Перевод на русский язык: Гудмэн И. Нечеткие множества как классы эквивалентности случайных множеств. - В сб.: Нечеткие множества и теория возможностей. Последние достижения. - М.: Радио и связь, 1986. - С. 241-264.)

10. Орлов А.И. Математика нечеткости. - Наука и жизнь. 1982. No.7. С.60-67.

Контрольные вопросы и задачи

1. В каких случаях целесообразно применение нечетких множеств?

2. Как с точки зрения нечетких множеств можно интерпретировать вероятность накрытия определенной точки случайным множеством?

3. Справедливо ли для нечетких множеств равенство $(A+B)C = AC + BC$? А равенство $(AB)C = (AC)(BC)$?

4. На множестве $Y = \{y_1,y_2,y_3\}$ задано нечеткое множество B с функцией принадлежности $\mu_B(y)$, причем $\mu_B(y_1) = 0{,}1$, $\mu_B(y_2) = 0{,}2$, $\mu_B(y_3) = 0{,}3$. Постройте случайное множество A так, чтобы Proj $A = B$.

5. На множестве $Y = \{y_1,y_2,y_3\}$ задано нечеткое множество B с функцией принадлежности $\mu_B(y)$, причем $\mu_B(y_1) = 0{,}2$, $\mu_B(y_2) = 0{,}1$, $\mu_B(y_3) = 0{,}5$. Постройте случайное множество A так, чтобы Proj $A = B$.

6. На множестве $Y = \{y_1,y_2,y_3\}$ задано нечеткое множество B с функцией принадлежности $\mu_B(y)$, причем $\mu_B(y_1) = 0{,}5$, $\mu_B(y_2) = 0{,}4$, $\mu_B(y_3) = 0{,}7$. Постройте случайное множество A так, чтобы Proj $A = B$.

7. На множестве $Y = \{y_1,y_2,y_3\}$ задано нечеткое множество B с функцией принадлежности $\mu_B(y)$, причем $\mu_B(y_1) = 0{,}3$, $\mu_B(y_2) = 0{,}2$, $\mu_B(y_3) = 0{,}1$. Постройте случайное множество A так, чтобы Proj $A = B$.

8. Опишите с помощью нечеткого подмножества временной шкалы понятие «молодой человек».

9. Опишите с помощью теории нечеткости понятие «куча зерен».

10. Как можно проводить кластерный анализ совокупности нечетких множеств?

Темы докладов, рефератов, исследовательских работ

1. Обсудите суждение: «Мы мыслим нечетко» (см. [10]). Почему нечеткость мышления помогает взаимопониманию?

2. Взаимосвязь теории нечеткости и теории вероятностей.

3. Методы оценивания функции принадлежности.

4. Теория нечеткости и интервальная математика.

5. Описание данных для выборок, элементы которых – нечеткие множества.

6. Регрессионный анализ нечетких переменных (согласно подходу [1]).

7. Непараметрические оценки плотности распределения вероятностей в пространстве нечетких множеств (согласно подходу [3]).