

# О связи между нечеткими и случайными множествами

Е.В. Копьева

Красноярский государственный университет  
e-mail: kgenyv@chat.ru

Одним из основоположников теории нечетких или, иначе, размытых множеств является известный американский математик Л.А. Заде [1], [2]. Понятие нечеткого множества в его теории предполагается в качестве средства математического моделирования неопределенных понятий, которыми оперирует человек при описании своих представлений о реальной мире, своих желаний, целей и т.п. Здесь учитывается возможность постепенного перехода от принадлежности к непринадлежности элемента множеству. Иными словами, элемент, вообще говоря, имеет степень принадлежности множеству, которая может являться промежуточной между полной принадлежностью и полной непринадлежностью.

Например, пусть  $B = \{\text{множество молодых людей}\}$ . Поскольку возраст начинается с нуля, то нижняя граница этого множества должна быть нулевой. Верхняя граница, допустим, будет 20 лет. Таким образом получили четкий интервал  $B = [0, 20]$ . Но у многих людей возникнет вопрос: почему на своем 20-ом дне рождения кто-то молодой, а на следующем уже немолодой? Поэтому для множества  $B$  нужен более естественный способ описания, который ослабит строгое разделение между молодым и немолодым. Значит, кроме строгой принадлежности или непринадлежности к множеству, введем более гибкие фразы, типа *хорошо, он/она принадлежит немного больше к множеству молодых* или *он/она, почти не принадлежит к множеству молодых*. Для большей наглядности множество  $B$  можно изобразить графически (рис.1).

Для работы используются статистика продаж книг, в которой каждая книга отождествляется со своим названием. Обозначим:

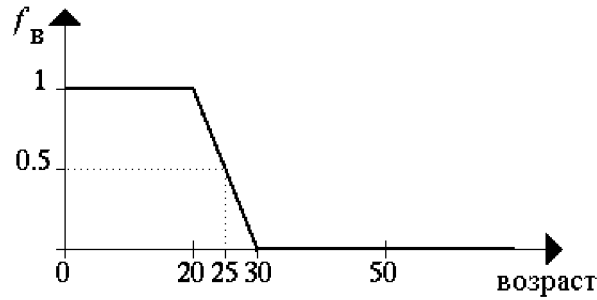


Рис. 1: Множество молодых людей, где  $f_B$  — функция степени принадлежности возраста человека к множеству В

- $\mathcal{R}$  — множество слов русского языка;
- $\mathfrak{B}$  — конечное множество (названий) книг (для простоты считаем, что каждое название  $b \in \mathfrak{B}$  — это некоторое подмножество различных слов русского языка  $b \subset \mathcal{R}$ );
- $\mathfrak{X}$  — конечное множество категорий, заранее выделенных в русском языке.

**Определение.** Под *категорией*  $x \in \mathfrak{X}$  понимается некоторое множество различных слов русского языка  $x \subset \mathcal{R}$ .

**Определение.** Говорят, что *книга* (название книги)  $b \in \mathfrak{B}$  *относится к категории*  $x \in \mathfrak{X}$  всякий раз, когда

$$b \cap x \neq \emptyset.$$

Так что каждая книга может относиться, вообще говоря, к нескольким категориям, или не относиться ни к одной. Поэтому каждой книге  $b \in \mathfrak{B}$  можно однозначно сопоставить подмножество категорий

$$X_b = \{x : x \cap b \neq \emptyset\} \subseteq \mathfrak{X}.$$

**Определение.** *Случайной книгой*  $K$  под множеством категорий  $\mathfrak{X}$  называется измеримое отображение

$$K : (\Omega, \mathcal{F}, \mathbf{P}) \rightarrow (2^{\mathfrak{X}}, 2^{2^{\mathfrak{X}}}),$$

значениями которого являются подмножества категорий.

Случайная книга определяется вероятностным распределением

$$p(X) = \mathbf{P}(K = X), \quad X \in 2^{\mathfrak{X}}$$

и имеет вероятности покрытия элементов вида

$$p_x = \mathbf{P}(x \in K), \quad x \in \mathfrak{X}.$$

**Определение.** *Нечеткой книгой*  $\mathcal{K}$  под множеством категорий  $\mathfrak{X}$  называется множество пар вида

$$\{(x, f_{\mathcal{K}}(x)), \quad x \in \mathfrak{X}\},$$

где  $f_{\mathcal{K}}(x)$  — функция степени принадлежности категории  $x$  нечеткой книге  $\mathcal{K}$ .

Рассмотрим операции пересечения и объединения над случайными книгами и над нечеткими книгами.

Для объединения двух случайных книг есть строгая формула для вероятности покрытия ею категории, которая зависит от совместного распределения первой и второй случайной книги и имеет вид:

$$P(x \in K_1 \cup K_2) = P(x \in K_1) + P(x \in K_2) - P(x \in K_1 \cap K_2),$$

где

$$P(x \in K_1 \cap K_2)$$

— вероятность покрытия категории  $x$  пересечением случайных множеств, которую можно оценить из статистики.

Другое дело, объединение двух нечетких книг. Теория нечетких множеств предлагает несколько произвольных формул для степени принадлежности категории  $x$  объединению двух нечетких книг. Например:

$$f(x \in K_1 \cup K_2) = \max\{f(x \in K_1), f(x \in K_2)\},$$

и т.д.

Так как вероятность покрытия и степень принадлежности теоретически совпадают у случайного и нечеткого множества, то и вероятности покрытия и степени принадлежности для их объединения должны совпадать. Поэтому необходимо посмотреть, существует ли связь

между функцией степени принадлежности нечеткого множества и вероятностями покрытия случайного множества.

В теории случайных множеств введено понятие ковариации [3], которая измеряет степень отклонения распределения случайного множества от независимого и которую можно определить следующим образом

$$\text{Kov}_X = p_X - \prod_{x \in X} p_x, \quad X \in 2^{\mathfrak{X}},$$

где  $p_X = \mathbf{P}(X \subseteq K)$ .

Тогда и возникла идея объяснить то разнообразие формул, которое существует в теории нечетких множеств для степеней принадлежности объединения двух нечетких множеств, тем, что всякий раз на практике специалистами по теории нечетких множеств подбирается такая степень принадлежности, которая как бы вбирает в себя статистическую зависимость между двумя случайными множествами, то есть, например, такая формула, что

$$f(x \in K_1 \cap K_2) \approx P(x \in K_1 \cap K_2).$$

Таким образом, разнообразие формул в теории нечетких множеств для вычисления степени принадлежности пересечения нечетких множеств объясняется тем, что иначе в этой теории нельзя описать зависимости между нечеткими множествами, аналоги (статистических зависимостей) которых для случайных описываются более естественно. Фактически в теории и практике нечётких множеств всякий раз решается одна главная задача отыскания функции  $C$  многих переменных<sup>1</sup>, которая связывает степень принадлежности пересечения нескольких нечетких множеств с их индивидуальными функциями принадлежности [3]:

$$f\left(x \in \bigcap_{i=1}^n K_i\right) = C(f(x \in K_1), \dots, f(x \in K_n)).$$

<sup>1</sup> Иногда подобные функции называют *копулой* (см., например, статью Е.В. Гришкиной в этой книге).

## Л И Т Е Р А Т У Р А

- [1] Орловский С.А. (1981) *Проблема принятия решений при нечеткой исходной информации*. М.: "Наука", Главная редакция физико-математической литературы, 208 с.
- [2] Заде Л.А. (1980) Размытые множества и их применение в распознавании образов и кластер-анализе. *Классификация и кластер*. М.: Мир, 208–247.
- [3] Воробьев О.Ю. (2002, в печати) *Теория случайных событий и её применение*. Красноярск: ИВМ СО РАН, 340 с.