Обобщение понятия границ Фреше

Аркадий Арсеньевич Новоселов
Сибирский федеральный университет
Институт математики и фундаментальной информатики
Красноярск
arcady@novosyolov.com

Аннотация. *В работе формулируется задача вычисления границ Фреше для произвольной теоретико-множественной операции, вводится понятие двойственных операций, изучается связь границ Фреше для двойственных операций, вводится понятие (ко)валентности событий, вычислены границы Фреше для всех операций на дуплете событий.*

Ключевые слова. *Событие, функция распределения, граница Фреше, терраска, валентность, двойственность, генератор.*

1 Введение

Первые упоминания о неравенствах типа границ Фреше появились, по-видимому, в работах Буля [1], а строгие формулировки и доказательства были даны Фреше в работах [2, 3].

В [4] описано понятие копулы и использованы границы Фреше для значений копулы (вероятности пересечения конечного набора событий).

Аналогичное понятие можно ввести и для других теоретико-множественных операций над n множествами.

Известно [5], что над n множествами можно определить $2^{(2^n)}$ различных теоретико-множественных операций. В настоящей работе обсуждается проблемы вычисления верхней и нижней границ Фреше для каждой такой операции.

В [6] была высказана идея о распространении понятия границ Фреше на произвольную теоретико-множественную операцию (объединение произвольного набора террасок, получаемых из заданного набора событий). В настоящей работе описывается задача вычисления границ Фреше для произвольной теоретико-множественной операции над n событиями, вводится понятие двойственных операций и изучается связь границ Фреше для таких операций, дано исчерпывающее описание границ Фреше для операций на дуплете событий.

Олег Воробьев (ред.), Труды XIII ФАМЭМС'2014, Красноярск: СФУ

2 Основные определения

Пусть $(\Omega, \mathcal{F}, \mathbf{P})$ — вероятностное пространство, зафиксируем целое положительное число n, обозначим $N = \{1, \dots, n\}$ и рассмотрим вектор $u \in [0,1]^n$, компоненты которого будем интерпретировать, как вероятности. Выберем n событий $x_1, \dots, x_n$ из σ-алгебры $\mathcal{F}$ таким образом, что $\mathbf{P}(x_i) = u_i$, $i \in N$, а в остальном произвольных. Какова максимальная и минимальная вероятность пересечения всех этих событий при сформулированных условиях? То есть, чему равны экстремальные значения

$$F_{\cap}^{-}(u) = \min \mathbf{P}\left(\bigcap_{i\in N} x_i\right), \tag{1}$$

$$F_{\cap}^{+}(u) = \max \mathbf{P}\left(\bigcap_{i\in N} x_i\right), \tag{2}$$

при условии

$$\mathbf{P}(x_i) = u_i,\ i \in N.$$

Ответ на этот вопрос дают классические границы Фреше. Когда вектор $(u_x, x \in \mathfrak{X})$ пробегает весь гиперкуб $[0,1]^n$, выражения (1), (2) задают на этом гиперкубе функции, называемые, соответственно, нижней и верхней границами Фреше для операции пересечения множеств.

Точные выражения границ Фреше для вероятности пересечения n событий тесно связаны с понятием копулы [4], поскольку значение совместной функции распределения случайного вектора $K = (K_1, \dots, K_n)$ совпадает с вероятностью пересечения событий $\{K_i \le u_i\}$, $i \in N$:

$$F_K(u) = \mathbf{P}\left(\bigcap_{i\in N}\{K_i \le u_i\}\right),$$

и выражается через функцию копулы $C : [0,1]^n \to [0,1]$ и маргинальные функции распределения F_i, $i \in N$ посредством

$$F_K(u) = C\left(F_1(u_1), \dots, F_n(u_n)\right).$$

Таким образом, функции (1), (2) являются верхними и нижними границами значений функций копулы C на гиперкубе $[0,1]^n$:

$$F_{\cap}^{-}(u) = \min_{u\in[0,1]^n} C(u),$$

$$F_{\cap}^{+}(u) = \max_{u \in [0,1]^n} C(u).$$

Границы для функции копулы известны и равны

$$F_{\cap}^{-}(u) = \max\left(0, \sum_{i \in N} u_i - (n-1)\right), \tag{3}$$

$$F_{\cap}^{+}(u) = \min(u_i,\, i \in N). \tag{4}$$

С помощью принципа дополнительности легко получить границы Фреше для операции объединения всех n событий. Действительно,

$$\bigcup_{i \in N} x_i = \left(\bigcap_{i \in N} x_i^c\right)^c,$$

так что

$$\begin{aligned}
\mathbf{P}\left(\bigcup_{i \in N} x_i\right) &= \left(1 - \mathbf{P}\left(\bigcap_{i \in N} x_i^c\right)\right) \\
&= 1 - \mathbf{P}\left(\bigcap_{i \in N} x_i^c\right) \geq 1 - \min_{i \in N}(1 - \mathbf{P}(x_i)) \\
&= \max_{i \in N} \mathbf{P}(x_i),
\end{aligned}$$

откуда нижняя граница для операции объединения всех событий записывается в виде функции

$$F_{\cup}^{-}(u) = \max_{i \in N} u_i. \tag{5}$$

С другой стороны,

$$\begin{aligned}
\mathbf{P}\left(\bigcup_{i \in N} x_i\right) &= 1 - \mathbf{P}\left(\bigcap_{i \in N} x_i^c\right) \\
&\leq 1 - \max\left(0, \sum_{i \in N}(1 - \mathbf{P}(x_i)) - (n-1)\right) \\
&= \min\left(1, \sum_{i \in N} \mathbf{P}(x_i)\right),
\end{aligned}$$

поэтому верхняя граница Фреше для объединения имеет вид

$$F_{\cup}^{+}(u) = \min\left(1, \sum_{i \in N} u_i\right). \tag{6}$$

3 Другие операции: подходы

3.1 Представление произвольной операции генератором

Множество из n событий разбивает Ω на 2^n террасок, каждой из которых может быть поставлено в соответствие множество наступивших событий. Обозначим множество номеров этих событий $I \subseteq N$, а соответствующую терраску

$$ter(I) = \left(\bigcap_{i \in I} x_i\right) \bigcap \left(\bigcap_{i \in N-I} x_i^c\right) \tag{7}$$

Так, например, терраске пересечения всех событий соответствует множество событий $\{x_1, \ldots, x_n\}$ или множество их индексов N. Другим террасскам соответствуют подмножества индексов $I \subseteq N$. А произвольная теоретико-множественная операция над n событиями представляется в виде объединения некоторого множества террасок, которому можно поставить в соответствие набор подмножеств индексов $A \subseteq 2^N$.

Рассмотрим эти понятия на примерах. Дуплет событий $\{x_1, x_2\}$ разбивает Ω на 4 терраски, каждая из которых соответствует одному их множеств событий $\emptyset, \{x_1\}, \{x_2\}, \{x_1, x_2\}$. Пересечение всех множеств представляется терраской $ter(\{x_1, x_2\})$, а объединение всех множеств — суммой террасок $ter(\{x_1\}) + ter(\{x_2\}) + ter(\{x_1, x_2\})$. Операции объединения соответствует набор террасок, задаваемых набором множеств индексов $A = \{\{1\}, \{2\}, \{1, 2\}\} \subseteq 2^N$. Будем называть множество террасок A, порождающее операцию, *генератором* этой теоретико-множественной операции.

Таким образом, для произвольного конечного множества событий произвольная теоретико-множественная операция над событиями из этого множества формально определяется своим генератором $A \subseteq 2^N$ посредством

$$\mathcal{O}_A = \sum_{I \in A} ter(I). \tag{8}$$

3.2 Границы Фреше для произвольной операции

Формализуем понятие границ Фреше для произвольной теоретико-множественной операции над событиями из конечного множества событий. Нижние и верхние границы Фреше для операции $\mathcal{O}_A$ будем обозначать, соответственно $F_A^{-}(u)$, $F_A^{+}(u)$, $u \in [0,1]^n$.

Зафиксируем число событий n, а также генератор A, представляющий операцию. Для каждой точки $u \in [0,1]^n$ и всевозможных наборов n событий из $\mathcal{F}$, удовлетворяющих условиям $\mathbf{P}(x_i) = u_i$, $i \in N$ вычислим наименьшее и наибольшее значение вероятности события $\mathbf{P}(\mathcal{O}_A(x_1, \ldots, x_n))$. Другими словами, нижняя граница Фреше является решением экстремальной задачи

$$F_A^{-}(u) = \min_{x_1, \ldots, x_n \in \mathcal{F}} \mathbf{P}(\mathcal{O}_A(x_1, \ldots, x_n)) \tag{9}$$

при условии

$$\mathbf{P}(x_i) = u_i,\, i \in N, \tag{10}$$

а верхняя граница Фреше - решением экстремальной задачи

$$F_A^{+}(u) = \max_{x_1, \ldots, x_n \in \mathcal{F}} \mathbf{P}(\mathcal{O}_A(x_1, \ldots, x_n)) \tag{11}$$

при том же условии (10).

Мощность генератора A будем называть *валентностью* операции $\mathcal{O}_A$: $v(\mathcal{O}_A) = |A|$; валентность может принимать значения от 0 до $2^n = 2^{|N|}$. Дополнение валентности до 2^n будем называть *ковалентностью* операции $\mathcal{O}_A$: $c(\mathcal{O}_A) = 2^{|N|} - |A|$.

Очевидно, $\mathcal{O}_\emptyset = \emptyset$ и $\mathcal{O}_{2^N} = \Omega$, поэтому

$$F^-_\emptyset(u) = F^+_\emptyset(u) = 0,\ F^-_{2^N}(u) = F^+_{2^N}(u) = 1, \qquad (12)$$

что исчерпывает вопрос о вычислении границ Фреше для операций нулевой валентности и полной валентности.

3.3 Двойственность

Элементарно выводятся следующие соотношения двойственности:

$$F^-_A(u) + F^+_{2^N-A}(u) = 1,\ F^+_A(u) + F^-_{2^N-A}(u) = 1. \qquad (13)$$

Вообще, назовем операции, соответствующие взаимно дополнительным генераторам A и $2^N - A$, *двойственными*. Ввиду соотношений двойственности (13) вычисление границ Фреше достаточно произвести только для половины операций, например, для операций небольшой валентности (не превышающей 2^{n-1}).

4 Границы Фреше для операций малой валентности

4.1 Одновалентная операция

Одноэлементный генератор $A = \{I\}$, где $I \subseteq N$ — некоторое подмножество индексов, порождает операцию, значением которой является терраска (7), поэтому нижняя граница Фреше имеет вид

$$\begin{aligned} & F^-_{\{I\}}(u) \\ = & \max\left(0, \sum_{i\in I} u_i + \sum_{i\in N-I}(1-u_i) - (n-1)\right) \\ = & \max\left(0, \sum_{i\in I} u_i - \sum_{i\in N-I} u_i - |I| + 1\right), \end{aligned} \qquad (14)$$

а верхняя граница Фреше —

$$F^+_{\{I\}}(u) = \min(\min_{i\in I} u_i, \min_{i\in N-I}(1-u_i)). \qquad (15)$$

Из соотношений двойственности (13) легко выводятся и границы Фреше для операций A ковалентности 1. В этом случае дополнение к A является одноэлементным, и границы получаются из (14), (15) в виде

$$\begin{aligned} & F^-_{2^N-\{I\}}(u) = 1 - F^+_{\{I\}}(u) \\ = & \max\left(\max_{i\in N}(1-u_i), \max_{i\in N-I} u_i\right), \end{aligned} \qquad (16)$$

$$\begin{aligned} & F^+_{2^N-\{I\}}(u) = 1 - F^-_{\{I\}}(u) \\ = & 1 - \max\left(0, \sum_{i\in I} u_i - \sum_{i\in N-I} u_i - |I| + 1\right) \\ = & \min\left(1, \sum_{i\in I}(1-u_i) + \sum_{i\in N-I} u_i\right) \end{aligned} \qquad (17)$$

4.2 Дуплет

Для дуплета событий $n = 2$, так что $\mathfrak{X} = \{x_1, x_2\}$. При этом валентность операции может изменяться от 0 до $2^2 = 4$. Ввиду соотношений двойственности достаточно вычислить границы Фреше для операций до валентности 2 включительно. Всего имеется одна нуль-валентная операция, $C^1_4 = 4$ одновалентных операций, $C^2_4 = 6$ двухвалентных операций, а также 4 трехвалентных операции и одна четырехвалентная.

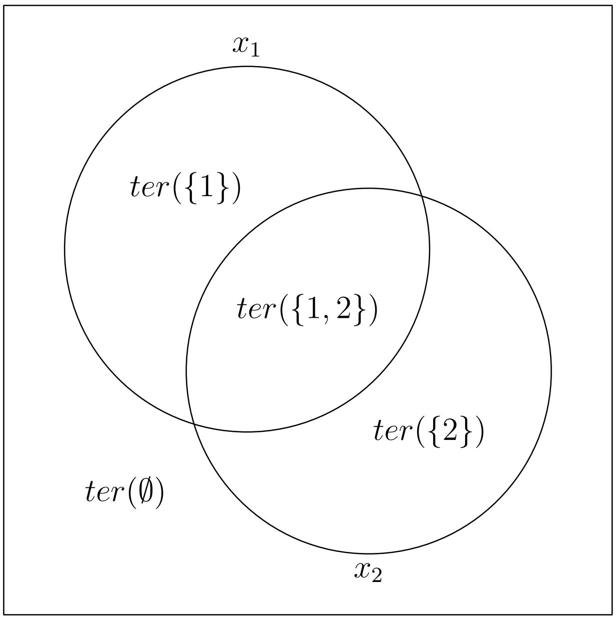

Рис. 1: Диаграмма Венна для дуплета событий $\{x_1, x_2\}$.

Для нуль-валентной и двойственной к ней четырехвалентной операции границы Фреше указаны в (12). Для одновалентных операций эта работа была проделана в параграфе 4.1, а выводы для трехвалентных операций получаются из соотношений двойственности (13), см. (16), (17). Здесь сразу перейдем к двухвалентным операциям, $|A| = 2$.

Шесть двухвалентных операций представлены в табл. 1 в виде двух групп (основные и двойственные)

Из таблицы видно, что результат операции для генератора $A = \{\{1\}, N\}$ не зависит от взаимного расположения событий x_1, x_2, и равен x_1, поэтому верхние и нижние границы Фреше легко вычисляются:

$$F^-_{\{\{1\},N\}}(u) = F^+_{\{\{1\},N\}}(u) = u_1. \qquad (18)$$

Аналогичный вывод справедлив для операции с ге-

A	$\mathcal{O}_A$	$2^N - A$	$\mathcal{O}_{2^N-A}$
$\{\{1\},N\}$	x_1	$\{\{2\},\emptyset\}$	x_1^c
$\{\{2\},N\}$	x_2	$\{\{1\},\emptyset\}$	x_2^c
$\{\{1\},\{2\}\}$	$x_1 \Delta x_2$	$\{\emptyset,N\}$	$ter(\emptyset)+ter(N)$

Таблица 1: Двухвалентные операции на дуплете и их результат

нератором $\{\{2\},N\}$: ее результат операции всегда равен x_2, а границы Фреше имеют вид

$$F^-_{\{\{2\},N\}}(u) = F^+_{\{\{2\},N\}}(u) = u_2. \tag{19}$$

Из соотношений двойственности (13) непосредственно выводим

$$F^-_{\{\{2\},\emptyset\}}(u) = F^+_{\{\{2\},\emptyset\}}(u) = 1 - u_1 \tag{20}$$

и

$$F^-_{\{\{1\},\emptyset\}}(u) = F^+_{\{\{1\},\emptyset\}}(u) = 1 - u_2. \tag{21}$$

Далее, для генератора $\{\{1\},\{2\}\}$ значением операции является симметрическая разность событий, что позволяет легко вычислить границы Фреше. Действительно, вероятность симметрической разности событий служит метрикой, наименьшее значение которой достигается при вложенном положении событий; метрика при этом равна модулю разности вероятностей событий, так что

$$F^-_{\{\{1\},\{2\}\}}(u) = |u_1 - u_2|. \tag{22}$$

Наибольшее значение расстояния между событиями достигается при максимальном их разделении, когда они либо не пересекаются, либо заполняют все Ω, что дает выражение для верхней границы Фреше

$$F^+_{\{\{1\},\{2\}\}}(u) = \min(1, u_1 + u_2). \tag{23}$$

Границы Фреше для генератора $\{\emptyset,\{N\}\}$ получаются из (22), (23) посредством соотношений двойственности (13):

$$\begin{aligned} F^-_{\{\emptyset,\{N\}\}}(u) &= \max(0, 1 - u_1 - u_2), \\ F^+_{\{\emptyset,\{N\}\}}(u) &= 1 - |u_1 - u_2|. \end{aligned}$$

4.3 Триплет

Для операций на триплете событий максимальная валентность составляет 8, и вычисление придется проводить до операций валентности 4 включительно. Эту работу еще предстоит проделать в будущем.

5 Заключение

В работе поставлена задача вычисления границ Фреше для произвольной теоретико-множественной операции. Для решения этой задачи предложено представление произвольной теоретико-множественной операции генератором, введены понятия двойственной операции, валентности и ковалентности операции. Проведено полное вычисление границ Фреше для операций на дуплете событий.

Благодарности

Автор признателен Олегу Юрьевичу Воробьеву за идею обобщения понятия границы Фреше на произвольные теоретико-множественные операции над n множествами (событиями).

Список литературы

[1] G. Boole. An investigation of the laws of thought, on which are founded the mathematical theories of logic and probability. 1854.

[2] M. Frechet. Generalisations du theoreme des probabilites totales. *Fundamenta Mathematica*, 25:379–387, 1935.

[3] M. Frechet. Sur les tableaux de correlation dont les marges sont donnees. *Annales de l'Universite de Lyon, Section A: Sciences mathematiques et astronomie*, 9:53–77, 1951.

[4] A. Sklar. Fonctions de repartition a n dimensions et leures marges. *Publications de l'Institut de Statistique de L'Universite de Paris*, 8:229–231, 1959.

[5] О. Ю. Воробьев. *Эвентология.* Сибирский федеральный университет, Красноярск, 2007, 435с., https://www.academia.edu/179393/Vorobyev_O.Yu._Eventology._-_Krasnoyarsk_Siberian_Federal_University._-_2007._-_435p.

[6] О. Ю. Воробьев. Границы Фреше для произвольной теоретико-множественной операции. Устное сообщение, 2011.