A generalization of Frechet bounds concept

Arcady Novosyolov
Siberian Federal University
Institute of Mathematics and Fundamental Informatics
Russia
arcady@novosyolov.com

Abstract. *A problem of Frechet bounds calculation for an arbitrary set theoretic operation is formulated and studied in the current paper. We introduce the concept of dual operations, study the relations between bounds for dual operations, define the valence (covalence) of operations, and calculate Frechet bounds for operations on two-element sets.*

Keywords. *Event, distribution function, Frechet bound, terrace, valence, duality, generator.*

1 Introduction

The first mentioning of inequalities resembling Frechet bounds perhaps may be found as long ago as in Bool's paper [1], though strict formulation and proof is due to Frechet, and was made in his papers [2, 3].

In [4] the concept of copula was introduces and Frechet bounds were established for copula function values (probability of intersection of a finite set of events).

A similar concept may be introduced for other set-theoretic operation on n sets.

It is known [5] that n sets allow as many as $2^{(2^n)}$ different set-theoretic operations. The present paper is devoted to the problem of calculation of upper and lower bounds for each operation of the sort.

In [6] the idea of generalizing the Frechet bounds concept onto any set-theoretic operation has been formulated. The present paper describes the problem of calculating Frechet bounds; for that matter we introduce the concept of dual operation, and the concept of operation's valency. We provide a complete description of Frechet bounds for a pair of events.

2 Basic definitions

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, fix a positive integer n, denote $N = \{1, \ldots, n\}$, and consider a vector $u \in [0,1]^n$ with components treated as probabilities. Select n events $x_1, \ldots, x_n$ in the σ-algebra $\mathcal{F}$ in such a way that $\mathbf{P}(x_i) = u_i$, $i \in N$. What is the minimum and maximum probability of intersection of all the events under given conditions? That is, what are the extreme values

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

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

given that

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

The answer is contained in the classic Frechet bounds. When the vector $(u_x, x \in \mathfrak{X})$ goes through the hypercube $[0,1]^n$, the expressions (1), (2) define two functions on the hypercube, which are called lower and upper Frechet bounds for sets intersection.

Exact values of Frechet bounds for intersection of n events are closely related to the copula concept [4], since the joint probability distribution function of the random vector $K = (K_1, \ldots, K_n)$ coincides with the probability of intersection of events $\{K_i \le u_i\}$, $i \in N$:

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

and is expressed via copula function $C : [0,1]^n \to [0,1]$ and the marginal distribution functions F_i, $i \in N$ as follows:

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

Thus the functions (1), (2) are lower and upper bounds of values of copula function C on the hypercube $[0,1]^n$:

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

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

The bounds for copula function are known and equal to

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

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

Oleg Vorobyev (ed.), Proc. XIII FAMEMS'2014, Krasnoyarsk: SFU

Using complementarity principle one can easily get Frechet bounds for the union of all n events. Indeed,

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

thus

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

and the lower bound for the union takes the form

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

On the other hand,

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

so the upper Frechet bound of the union is

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

3 Other operations: some approaches

3.1 Representation of a set operation by generator

A set of n events splits Ω up to 2^n terraces, each of which may de described by the set of events occurred. Denote the set of numbers of these events by $I \subseteq N$, and the corresponding terrace by

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

E.g. intersection of all events is the terrace corresponding to the set of events $\{x_1,\dots,x_n\}$ or the set of their indices N. Other terraces correspond to other index subsets $I \subseteq N$. Any set operation with n events is a union of a number of terraces, which is represented by a collection on index subsets $A \subseteq 2^N$.

Consider examples. A pair of events $\{x_1,x_2\}$ splits Ω to 4 terraces, corresponding to one of the events sets $\emptyset, \{x_1\}, \{x_2\}, \{x_1,x_2\}$. Intersection of all sets is represented by the terrace $ter(\{x_1,x_2\})$, and union of all sets is represented by the sum of terraces $ter(\{x_1\}) + ter(\{x_2\}) + ter(\{x_1,x_2\})$. Union operation relates to the collection of terraces $A = \{\{1\},\{2\},\{1,2\}\} \subseteq 2^N$. We will call *generator* of operation the set A defining that operation.

Thus for any finite set of events an arbitrary operation on this set is defined by its generator A by

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

3.2 Frechet bounds for a set operation

Let us formalize the concept of Frechet bounds for an operation over events from a finite set of events. Denote $F_A^-(u)$, $F_A^+(u)$, $u \in [0,1]^n$ lower and upper bounds for an operation $\mathcal{O}_A$.

Next let us fix the events number n and generator A for the operation. For each point $u \in [0,1]^n$ and any collection of n events in $\mathcal{F}$, satisfying $\mathbf{P}(x_i) = u_i$, $i \in N$ calculate the minimum and maximum values of the probability $\mathbf{P}(\mathcal{O}_A(x_1,\dots,x_n))$. In other words, lower Frechet bound is a solution for the extreme problem

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

subject to

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

and the upper Frechet bound is a solution of the extreme problem

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

subject to the same condition (10).

We will call *valency* of the operation $\mathcal{O}_A$ the power of its generator A: $v(\mathcal{O}_A) = |A|$; valency may take values from 0 to $2^n = 2^{|N|}$. Complement of the valency to 2^n is *covalency* of the operation $\mathcal{O}_A$: $c(\mathcal{O}_A) = 2^{|N|} - |A|$.

Clearly $\mathcal{O}_\emptyset = \emptyset$ and $\mathcal{O}_{2^N} = \Omega$, thus

$$F_\emptyset^-(u) = F_\emptyset^+(u) = 0,\ F_{2^N}^-(u) = F_{2^N}^+(u) = 1, \tag{12}$$

which gives the complete solution of Frechet bounds problem for operations of zero and full valency.

3.3 Duality

The following duality relations are straightforward:

$$F_A^-(u) + F_{2^N-A}^+(u) = 1,\ F_A^+(u) + F_{2^N-A}^-(u) = 1. \tag{13}$$

Generally speaking, we will call operation *dual*, if they correspond to mutually complementary generators A and $2^N - A$. Because of duality relations (13) we need to calculate Frechet bounds only for a half of all operations, say, for low valency operations, such that $v(\mathcal{O}) \leq 2^{n-1}$.

4 Frechet bounds for low valency operations

4.1 Univalent operation

Single-element generator $A = \{I\}$, where $I \subseteq N$ is an index subset, generates an operation with a terrace as a value (7), so the lower Frechet bound has the form

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

and the upper Frechet bound looks like

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

Duality relations (13) easily imply Frechet bounds for operations A of covalency 1. In this case complement of A has only one element, thus bounds follows from (14), (15) as

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

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

4.2 Duplet

For a duplet of events we have $n = 2$, thus $\mathfrak{X} = \{x_1, x_2\}$. Operation valency ranges from 0 to $2^2 = 4$. Keeping duality relations in mind, it is sufficient to calculate Frechet bounds for operations with valency up to 2. There is one zero-valent operation, $C_4^1 = 4$ univalent operations, $C_4^2 = 6$ bivalent operations, and also 4 trivalent operations and one 4-valent operation.

Frechet bounds for zero-valency and four-valency operations are presented in (12). Univalent operations were covered in section 4.1, and derivation for 3-valent operations follows from duality relations (13), see (16), (17). So we would move to bivalent operations, $|A| = 2$.

The six bivalent operations are presented in the Table 1 in two groups: basic and dual operations.

We see from the table that the operation result for the generator $A = \{\{1\}, N\}$ does not depend on x_2, and equals x_1, so the Frechat bounds are easy:

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

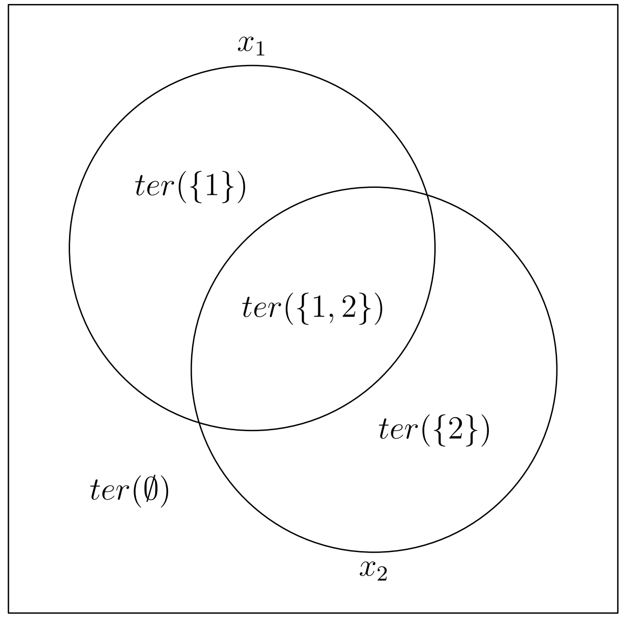

Figure 1: Venn diagram for an events duplet $\{x_1, x_2\}$.

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
$\{\{1\}, \{2\}\}$	$x_1 \Delta x_2$	$\{\emptyset, N\}$	$ter(\emptyset) + ter(N)$

Table 1: Bivalent operations on a duplet and their results

Similar derivation applies to the operation with generator $\{\{2\}, N\}$: its result always equals x_2, and Frechet bounds look like

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

Duality relations (13)imply

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

and

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

Next, for the generator $\{\{1\}, \{2\}\}$, the operation value is the symmetric difference of events, so the Frechet bounds are easy. Indeed, probability of symmetric difference is a metric, reaching its minimum value for included events, the latter being equal absolute value of the difference of probabilities of events, so

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

The largest metric value is attained when the events are extremely separated, that is, they are either do not intersect, or fill the whole Ω, which implies for the upper Frechet bound

$$\begin{aligned} F^+_{\{\{1\},\{2\}\}}(u) &= \min(u_1 + u_2, 2 - (u_1 + u_2)) \\ &= 1 - |u_1 + u_2 - 1|. \end{aligned} \tag{23}$$

Frechet bounds for the generator $\{\emptyset, \{N\}\}$ follow from (22), (23) using duality relations (13):

$$F^{-}_{\{\emptyset,\{N\}\}}(u) = |u_1 + u_2 - 1|,$$
$$F^{+}_{\{\emptyset,\{N\}\}}(u) = 1 - |u_1 - u_2|.$$

4.3 Triplet

For the operation of a triplet of events the maximum valency is 8, so we need to calculate Frechet bounds for operations with valency up to 4. This is a matter of future work.

5 Conclusion

We posed a problem of calculating Frechet bounds for arbitrary set-theoretic operation. To solve the problem we represent operation by its generator, introduce the concept of dual operation, valency and covalency of operations. We also proposed the complete solution for a duplet of operations.

Acknowledgements

The author is greatful to Oleg Vorobyov for multiple discussions of problems of random sets of events, and to Ludmila Shangareeva for her stimulating questions on measurability of random sets of events.

References

[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 n dimensions et leures marges. *Publications de l'Institut de Statistique de L'Universite de Paris*, 8:229–231, 1959.

[5] O. Yu. Vorobyev. *Eventology.* Siberian Federal University, Krasnoyarsk, 2007, 435p., (in Russian).

[6] O. Yu. Vorobyev. Frechet bounds for and arbitrary sret-theoretic operation. Private communication, 2011.