Глава 2

Теория свидетельств Демпстера–Шейфера

Тщательнее всего следует проверять правильность суждений, которые кажутся нам очевидными.

Марк Твен

2.1. Основные определения теории Демпстера–Шейфера

Теория Демпстера–Шейфера[1], также называемая математической *теорией свидетельств*[2], *теорией функций доверия*[3] или *теорией случайных множеств*[4], была предложена в работе Демпстера [24] и позднее развита в работе Шейфера [75] в качестве инструмента для моделирования и обработки неточных (интервальных) экспертных оценок, измерений или наблюдений.

Теория Демпстера–Шейфера использует математические объекты, называемые “функциями доверия”. Обычно их основная цель заключается в моделировании степени доверия некоторого субъекта к чему-либо. В то же время в литературе имеется большое количество интерпретаций функций доверия, которые могут использоваться в различных прикладных задачах. Как заметил Сметс [77], работа Демпстера [24] была

---

[1]Dempster-Shafer theory (англ.)
[2]Mathematical theory of evidence (англ.)
[3]Belief function theory (англ.)
[4]Theory of random sets (англ.)

посвящена многозначным отображениям и интервальным вероятностям и не была ориентирована в направлении моделирования субъективной степени доверия. Основной идеей Шейфера было использование математического аппарата, предложенного Демпстером, для того чтобы моделировать субъективную степень доверия. Однако Шейфер не доказывает и не обосновывает в своей основной книге [75] правомерность использования соответствующего математического аппарата. Впоследствии основная часть обоснований была сделана с использованием таких понятий, как случайные множества, многозначные отображения, нижние и верхние вероятности, которые в большей степени свойственны теории случайных множеств. Поэтому в дальнейшем будем рассматривать функции доверия в основном с позиции теории случайных множеств.

Случайные множества – это множества, полученные случайно, т.е., говоря более или менее строго, случайное множество – это случайная величина, принимающая в качестве значений некоторые множества вместо точек.

Пусть Ω – некоторое множество, которое в теории свидетельств иногда называется *универсальным множеством* или *фреймом различений*[5]. Предположим, что N наблюдений или измерений элемента $\omega \in \Omega$ было получено в качестве информации об объекте, принимающем значения из Ω. При этом предполагается, что результат измерений или наблюдений является неточным, т.е. представляет из себя некоторый интервал (подмножество) A значений Ω. Пусть c_i означает количество наблюдаемых подмножеств $A_i \subseteq \Omega$, а $\mathcal{P}o(\Omega)$ – множество всех подмножеств Ω. Частотная функция m, называемая *базовой вероятностью*[6], определяется как [24, 54, 75]:

$$m : \mathcal{P}o(\Omega) \to [0,1], \quad m(\varnothing) = 0, \quad \sum_{A \in \mathcal{P}o(\Omega)} m(A) = 1.$$

Заметим, что область определения базовых вероятностей $\mathcal{P}o(\Omega)$ отличается от области определения Ω функции распределения вероятностей. Согласно [24], базовая вероятность

[5] frame of discernment (англ.)
[6] basic probability assignment (англ.)

может быть получена следующим образом:

$$m(A_i) = c_i/N. \tag{2.1}$$

Если $m(A_i) > 0$, т.е. подмножество A_i в качестве результата измерения или наблюдения было получено хотя бы один раз, то A_i называется *фокальным элементом*[7].

Определим теперь функции доверия и правдоподобия. Согласно [75], *функция доверия*[8], обозначаемая $\mathrm{Bel}(A)$, и *функция правдоподобия*[9], обозначаемая $\mathrm{Pl}(A)$, события $A \subseteq \mathcal{P}o(\Omega)$ определяются как

$$\mathrm{Bel}(A) = \sum_{A_i : A_i \subseteq A} m(A_i), \quad \mathrm{Pl}(A) = \sum_{A_i : A_i \cap A \neq \varnothing} m(A_i). \tag{2.2}$$

Если известны функции доверия событий A_i, то базовые вероятности события A могут быть вычислены при помощи формулы инверсии Мёбиуса

$$m(A) = \sum_{A_i : A_i \subseteq A} (-1)^{|A - A_i|} \mathrm{Bel}(A_i),$$

где $|A|$ – мощность множества A (количество элементов множества A).

Как показано в работах [46, 47], функция доверия может быть формально определена как функция, удовлетворяющая аксиомам, являющимся ослабленным вариантом аксиом Колмогорова, характеризующих вероятность. Поэтому в некоторых случаях имеет смысл рассматривать функцию доверия (правдоподобия) как обобщенную вероятность, а доверие $\mathrm{Bel}(A)$ и правдоподобие $\mathrm{Pl}(A)$ как нижнюю и верхнюю вероятности события A, т.е. $\mathrm{Bel}(A) \leq \Pr(A) \leq \mathrm{Pl}(A)$.

Пример 2.1. Имеются четыре кандидата на некоторую должность ($\Omega = \{1, 2, 3, 4\}$). После интервью 10 экспертов ($N = 10$) пытаются определить наиболее вероятного кандидата. 5 экспертов ($c_1 = 5$) считают, что первый кандидат ($A_1 = \{1\}$) является наиболее предпочтительным, 2 эксперта ($c_2 = 2$) считают, что первый или

[7]focal element (англ.)

[8]belief function (англ.)

[9]plausibility function (англ.)

второй кандидат ($A_2 = \{1, 2\}$) является наиболее предпочтительным, 3 эксперта ($c_3 = 3$) выбирают третьего кандидата ($A_3 = \{3\}$). Используя (2.1), получим:

$$m(A_1) = c_1/N = 0.5,\ m(A_2) = c_2/N = 0.2,$$
$$m(A_3) = c_3/N = 0.3.$$

Найдем нижнюю и верхнюю границы для вероятности первого кандидата ($A = \{1\}$). Используя (2.2), получаем:

$$\mathrm{Bel}(A) = m(A_1) = 0.5,\ \mathrm{Pl}(A) = m(A_1) + m(A_2) = 0.7.$$

Нижняя и верхняя границы для вероятности второго кандидата ($A = \{2\}$):

$$\mathrm{Bel}(A) = 0,\ \mathrm{Pl}(A) = m(A_2) = 0.2.$$

Нижняя и верхняя границы для вероятности третьего кандидата ($A = \{3\}$):

$$\mathrm{Bel}(A) = m(A_3) = 0.3,\ \mathrm{Pl}(A) = m(A_3) = 0.3.$$

Функции доверия и правдоподобия могут быть определены также в случае, если множество Ω не является конечным, например Ω может быть множеством всех вещественных чисел. В этом случае расчеты остаются такими же, как это видно из следующего примера.

Пример 2.2. Опрос экспертов по поводу будущей цены акций некоторого предприятия дал следующие результаты: 4 экперта ($c_1 = 4$) предоставили интервал $A_1 = [30, 36]$, 1 эксперт ($c_2 = 1$) – интервал $A_2 = [28, 40]$, 5 экспертов ($c_3 = 5$) предоставили интервал $A_3 = [34, 38]$. Определим функцию доверия и правдоподобия интервала $A = [28, 32]$. Так как $N = c_1 + c_2 + c_3 = 10$, то базовые вероятности каждого интервала имеют следующий вид:

$$m(A_1) = c_1/N = 0.4,\ m(A_2) = c_2/N = 0.1,$$
$$m(A_3) = c_3/N = 0.5.$$

Тогда

$$\mathrm{Bel}(A) = 0,\ \mathrm{Pl}(A) = m(A_1) + m(A_2) = 0.5.$$

Следует отметить, что определение (2.1) может быть использовано, когда количество наблюдений N достаточно большое. Однако это условие может часто нарушаться. Если N является малым, то результаты расчетов получаются слишком рискованными, чтобы им доверять.

2.2. Функции доверия и множество полиномиальных моделей

Пусть $\Omega = \{\omega_1, ..., \omega_K\}$ – множество возможных исходов ω_j. Совокупность N наблюдений, независимо выбранных из Ω с одинаковыми вероятностями каждого исхода $\Pr\{\omega_j\} = \theta_j$ для всех $j = 1, ..., K$, где $\theta_j \geq 0$ и $\sum_{j=1}^{K} \theta_j = 1$, составляет стандартную полиномиальную модель. Обозначим $\theta = (\theta_1, ..., \theta_K)$. Пусть n_j – число наблюдений исхода ω_j в N опытах, таких что $n_j \geq 0$ и $\sum_{j=1}^{K} n_j = N$. При введенных предположениях случайные величины $n_1, ..., n_K$ имеют полиномиальное распределение, и функция правдоподобия, построенная на основе наблюдений $\mathbf{n} = (n_1, ..., n_K)$, имеет вид

$$L(\mathbf{n}|\theta) \propto \prod_{j=1}^{K} \theta_j^{n_j}.$$

Теперь предположим, что имеется множество N неточных наблюдений $A_i \subseteq \Omega$ и c_i – число появлений подмножества A_i, $i = 1, ..., n$. Базовые вероятности подмножеств A_i вычисляются по формуле (2.1). Оказывается, что на основе данной информации можно построить конечное множество полиномиальных моделей.

Для построения множества полиномиальных моделей рассмотрим урновую модель. Предположим, что имеется K урн, $\omega_1, ..., \omega_K$, содержащих шары с номерами $1, ..., K$ соответственно. Сначала случайным образом выбирается подмножество урн A_i, состоящее из l_i урн, такое что $A_i = \{\omega_j : j \in J_i\}$, т.е. подмножество урн A_i содержит шары с номерами из индексного множества J_i. Затем случайным образом вытаскиваются c_i шаров из урн с номерами из J_i. При этом нет никаких предпочтений по выбору шаров из урн, соответствующих A_i. Данная процедура повторяется n раз, т.е. для каждого фокального элемента. Что можно сказать о числе шаров с определенными номерами, вытащенных из урн? Очевидно, что не существует однозначного ответа, за исключением случая, когда все подмножества A_i состоят из одного элемента, так как выбор шара из урн, соответствующих подмножеству A_i, производится случайно. Таким образом, существует множество различных комбинаций номеров шаров. Предположим, что число таких комбинаций равно M, и обозначим k-й возможный вектор количества выбранных номеров шаров

$$\mathbf{n}^{(k)} = (n_1^{(k)}, ..., n_K^{(k)}), k = 1, ..., M,$$

где $n_j^{(k)}$ – количество случайно выбранных шаров с номером j. Кроме того, обозначим вектор числа появлений c_i подмножества A_i как $\mathbf{c} = (c_1, ..., c_n)$. Если предположить, что все подмножества A_i выбраны независимо из множества $\mathcal{P}o(\Omega)$ всех подмножеств Ω и вероятность выбора шара из j-й урны равна θ_j, то каждая комбинация шаров $\mathbf{n}^{(k)}$ образует стандартную полиномиальную модель. M возможных комбинаций шаров образуют M полиномиальных моделей. При этом все модели равноправны в том смысле, что нельзя выбрать одну из них, как наиболее предпочтительную по отношению к другим моделям.

Пример 2.3. Предположим, что $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, имеют место три фокальных элемента $A_1 = \{\omega_1\}$, $A_2 = \{\omega_1, \omega_2\}$, $A_3 = \{\omega_2, \omega_3\}$ и $c_1 = c_3 = 1$, $c_2 = 2$. Возможными комбинациями $\mathbf{n}^{(k)}$ числа шаров являются (3100), (2200), (1300), (3010), (2110), (1210).

Так как существует множество векторов $\mathbf{n}^{(k)}$, то существует и множество различных функций правдоподобия $L(\mathbf{n}^{(k)}|\theta)$, $k = 1, ..., M$. Отсюда следует, что вероятность произвольного события $A \subseteq \Omega$ зависит от $\mathbf{n}^{(k)}$, т.е. можно найти $P(A|\mathbf{n}^{(k)})$. Даже если бы условные вероятности $P(A|\mathbf{n}^{(k)})$ для каждого события A в Ω и для каждого вектора $\mathbf{n}^{(k)}$ были бы точно известны, можно было бы найти только нижнюю и верхнюю границы вероятности события A, которые определяются как

$$\underline{P}(A|\mathbf{c}) = \min_{k=1,...,M} P(A|\mathbf{n}^{(k)}), \quad \overline{P}(A|\mathbf{c}) = \max_{k=1,...,M} P(A|\mathbf{n}^{(k)}).$$

В частном случае, когда все подмножества A_i содержат только один элемент, $M = 1$ и

$$\underline{P}(A|\mathbf{c}) = P(A|\mathbf{n}^{(k)}), \quad \overline{P}(A|\mathbf{c}) = P(A|\mathbf{n}^{(k)}).$$

Так как векторы $\mathbf{n}^{(k)}$ зависят от $\mathbf{c}$, то результирующие нижняя и верхняя вероятности (после минимизации и максимизации $P(A|\mathbf{n}^{(k)})$) зависят от $\mathbf{c}$ и будут обозначаться $\underline{P}(A|\mathbf{c})$ и $\overline{P}(A|\mathbf{c})$.

Следует отметить, что все сказанное выше справедливо и для случая, когда Ω – некоторый интервал вещественных чисел. В этом случае можно преобразовать множество Ω с бесконечным числом элементов к конечному множеству. Действительно, пусть $\{\mathbf{i}\} = \{(i_1, ..., i_n, i_{n+1})\}$ – множество всех двоичных векторов, состоящих из $n+1$ элементов, таких что $i_j \in \{0,1\}$. Для каждого вектора $\mathbf{i}$ определим интервал B_k $(k = 1, ..., 2^{n+1})$ следующим образом:

$$B_k = \left(\bigcap_{j:i_j=1} A_j \right) \bigcap \left(\bigcap_{j:i_j=0} A_j^c \right), \ i_j \in \mathbf{i}.$$

Здесь $A_{n+1} = \Omega$. Таким образом, множество Ω разделено на конечное множество таких непересекающихся интервалов B_k, что $B_1 \cup ... \cup B_K = \Omega$, $K = 2^{n+1}$. Более того, каждый интервал A_i может быть представлен как объединение конечного числа интервалов B_k. Поэтому, ассоциируя каждый интервал B_k с урной, можно построить множество полиномиальных моделей, рассмотренных выше.

Заметим, что оценкой вероятности θ_i выбора шара из i-й урны является отношение $n_i^{(k)}/N$. При этом данная оценка зависит от номера возможной реализации k. Пусть J – множество индексов элементов A, т.е. $A = \{\omega_j : j \in J\}$. Тогда вероятность $P(A|\mathbf{n}^{(k)})$ события A при условии имеющихся данных $\mathbf{n}^{(k)}$ равна $\sum_{j \in J} n_j^{(k)}/N$. Отсюда можно найти

$$\underline{P}(A|\mathbf{c}) = \min_{k=1,...,M} \sum_{j \in J} n_j^{(k)}/N, \ \overline{P}(A|\mathbf{c}) = \max_{k=1,...,M} \sum_{j \in J} n_j^{(k)}/N.$$

Необходимо отметить, что количество шаров из A не может быть меньше, чем сумма всех шаров из подмножеств $A_i \subseteq A$, так как при выполнении условия $A_i \subsetneq A$ шары из A_i могут быть выбраны также из урн, находящихся вне A. Поэтому минимум суммы $\sum_{j\in J} n_j^{(k)}$ достигается для комбинации шаров $\mathbf{n}^{(k)}$, выбранных из подмножеств $A_i \subseteq A$, и

$$\min_{k=1,\dots,M} \sum_{j\in J} n_j^{(k)} = \sum_{i:A_i\subseteq A} c_i.$$

Количество шаров из A не может быть больше, чем сумма всех шаров из таких подмножеств A_i, что $A_i \cap A \neq \varnothing$. Следовательно, справедливо равенство

$$\max_{k=1,\dots,M} \sum_{j\in J} n_j^{(k)} = \sum_{i:A_i\cap A\neq\varnothing} c_i.$$

Отсюда

$$\underline{P}(A|\mathbf{c}) = \sum_{i:A_i\subseteq A} c_i/N = \sum_{i:A_i\subseteq A} m(A_i) = \mathrm{Bel}(A),$$

$$\overline{P}(A|\mathbf{c}) = \sum_{i:A_i\cap A\neq\varnothing} c_i/N = \sum_{i:A_i\cap A\neq\varnothing} m(A_i) = \mathrm{Pl}(A).$$

Таким образом, получены функции доверия и правдоподобия, как нижняя и верхняя вероятности события A, на основе модифицированной схемы урн и шаров.

2.3. Функции доверия и случайные множества

Функцию доверия и функцию правдоподобия можно получить также в терминах многозначного отображения или случайных множеств. Рассмотрим вероятности $P(\psi)$, определенные на множестве Ψ (это множество наших наблюдений), которое связано с множеством Ω (это множество возможных значений наших наблюдений) через многозначное отображение $G : \Psi \to \mathcal{P}o(\Omega)$ (см. рис. 2.1). Тогда базовые вероятности вычисляются следующим образом [24]:

$$m(A_i) = P(\psi_i) = c_i/N,\ \psi_i \in \Psi.$$

Это многозначное отображение выражает неточность измерения (наблюдения), которая определяется в процессе этих измерений (наблюдений). Другими словами, многозначное

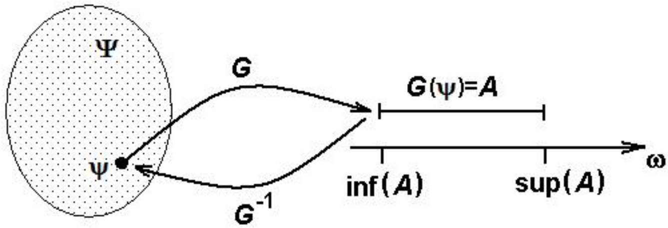

Рис. 2.1. Пример многозначного отображения

отображение выражает нашу неспособность приписать отдельное число каждому наблюдению. Для каждого множества $A_i \in \mathcal{P}o(\Omega)$ значение базовой вероятности $m(A_i)$ является вероятностью точки $\psi_i = G^{-1}(A_i)$ $(\psi_i \in \Psi)$. *Случайное множество*[10] есть пара $(\mathcal{F}, m)$, где $\mathcal{F}$ – семейство всех N фокальных элементов. Если множество Ω является декартовым произведением k множеств, т.е. $\Omega = \Omega_1 \times ... \times \Omega_k$, то вместо случайного множества имеем *случайное отношение*[11].

Таким образом, в теории случайных множеств вместо подсчета появлений отдельных элементов множества Ω подсчитываются наблюдения подмножеств $A_i \subseteq \Omega$ и каждое наблюдение подмножества $A_i \subseteq \Omega$ указывает, что событие имеет место где-то в A_i без конкретного определения вероятностей в каждой точке A_i. Любое распределение вероятностей с суммарным весом $m(A_i)$ может быть определено на интервале (подмножестве) A_i, и все они равноправны в том смысле, что ни одно из них не может быть выбрано как наиболее предпочтительное без дополнительной информации.

Пусть A – произвольное подмножество множества Ω, т.е. $A \subseteq \Omega$. Если определить X_* как подмножество Ψ, элементы которого *должны обязательно* вести через многозначное отображение G к A, т.е.

$$X_* = \{\psi \in \Psi : G(\psi) \subseteq A\},$$

то нижняя граница вероятности подмножества A, в соответ-

[10]random set (англ.)
[11]random relation (англ.)

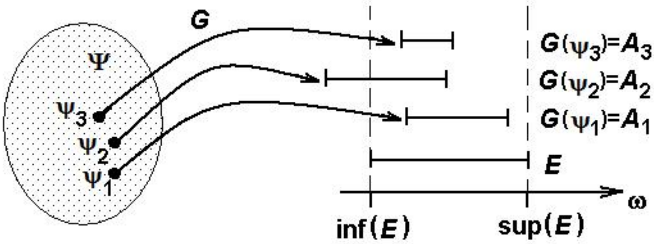

Рис. 2.2. Соотношение различных интервалов при многозначном отображении

ствием с принципом индуктивного вывода Демпстера [24], определяется как $\underline{P}(A|\mathbf{c}) = P(X_*)$. При этом

$$P(X_*) = \sum_{i:\psi_i \in X_*} P(\psi_i) = \sum_{i:G(\psi_i)\subseteq A} m(A_i) = \sum_{i:A_i \subseteq A} m(A_i).$$

Если определить X^* как подмножество Ψ, элементы которого *могут* вести через многозначное отображение G к A, т.е.

$$X^* = \{\psi \in \Psi : G(\psi) \cap A \neq \varnothing\},$$

то верхняя граница вероятности подмножества A определяется как $\overline{P}(A|\mathbf{c}) = P(X^*)$. При этом

$$P(X^*) = \sum_{i:\psi_i \in X^*} P(\psi_i) = \sum_{i:G(\psi_i)\cap A \neq \varnothing} m(A_i) = \\ = \sum_{i:A_i \cap A \neq \varnothing} m(A_i).$$

Таким образом, для произвольного подмножества $E \subset \Omega$ можно определить только нижнюю и верхнюю границы его вероятности. При этом нижняя граница $\underline{P}(E|\mathbf{c})$ есть функция доверия $\mathrm{Bel}(E)$, а верхняя граница $\overline{P}(E|\mathbf{c})$ – функция правдоподобия $\mathrm{Pl}(E)$.

Пример 2.4. В качестве иллюстрации сказанного рассмотрим пример многозначного отображения, показанного на рис. 2.2. Из рис. 2.2 видно, что имеют место три интервала наблюдений A_1, A_2 и A_3, которым соответствуют три точки на множестве Ψ. Таким образом, подмножеству E на Ω соответствует подмножество $X_* = \{\psi_1, \psi_3\}$. Следовательно,

$$\mathrm{Bel}(E) = \Pr(\{\psi_1, \psi_3\}) = P(\psi_1) + P(\psi_3) = m(A_1) + m(A_3).$$

Из рис. 2.2 также следует, что подмножеству E соответствует подмножество $X^* = \{\psi_1, \psi_2, \psi_3\}$. Отсюда

$$\begin{aligned}\mathrm{Pl}(E) &= \Pr(\{\psi_1, \psi_2, \psi_3\}) = P(\psi_1) + P(\psi_2) + P(\psi_3) = \\ &= m(A_1) + m(A_2) + m(A_3).\end{aligned}$$

2.4. Нижняя и верхняя функции распределения

Предположим, что множество Ω является множеством или подмножеством всех действительных чисел, т.е. $\Omega = \mathbb{R}$ или $\Omega \subset \mathbb{R}$. Кроме того, имеется множество наблюдаемых интервалов (фокальных элементов) A_i, $i = 1, ..., n$, с базовыми вероятностями, определенными в (2.1). Если говорить о функциях доверия и правдоподобия как о нижней и верхней границах вероятности, то, рассматривая интервалы $E_x = (-\infty, x]$, можно построить нижнюю и верхнюю функции распределения вероятностей случайной величины, информация о которой дана в виде набора фокальных элементов A_i совместно с ненулевыми базовыми вероятностями $m(A_i)$.

Пусть Ω – подмножество действительных чисел, ограниченных значениями Ω_* и Ω^*. Тогда нижняя $\underline{F}$ и верхняя $\overline{F}$ функции распределения вероятностей случайной величины X, о которой имеются данные в виде интервалов A_i, $i = 1, ..., n$, принимают вид

$$\underline{F}(x) = \underline{P}(\{\omega \le x\}) = \begin{cases} \sum_{i:\sup A_i \le x} c_i/N, & x < \Omega^* \\ 1, & x = \Omega^* \end{cases},$$

$$\overline{F}(x) = \overline{P}(\{\omega \le x\}) = \begin{cases} \sum_{i:\inf A_i \le x} c_i/N, & x > \Omega_* \\ 0, & x = \Omega_* \end{cases}.$$

Эти функции распределения являются границами для всех возможных функций распределения, которые совместимы с имеющимися данными. Кроме того, получаемые границы являются ступенчатыми и могут рассматриваться в рамках аппарата “вероятностных блоков” или “p-блоков”[12]. P-блоки[13] [13] – это класс всех распределений вероятностей, ограниченных некоторыми нижними и верхними функциями распределения. Как показано в работах [13, 60], каждая структура Демпстера–Шейфера однозначно определяет уникальную структуру p-блоков и каждая структура p-блоков имеет эквивалентный набор структур Демпстера–Шейфера.

Так как полученные функции распределения являются границами для всех возможных функций распределения случайной величины X, то можно вычислить нижнее и верхнее математические ожидания случайной величины X. При этом границы математического ожидания полностью определяются только граничными функциями распределения. Если X – непрерывная случайная величина, то

$$\underline{\mathbb{E}}X = \int_\Omega \omega \frac{\mathrm{d}\overline{F}(x)}{\mathrm{d}x}\mathrm{d}\omega = N^{-1}\sum_{i=1}^{n} c_i \cdot \inf A_i = \\ = \sum_{i=1}^{n} m(A_i)\cdot \inf A_i, \tag{2.3}$$

$$\overline{\mathbb{E}}X = \int_\Omega \omega \frac{\mathrm{d}\underline{F}(x)}{\mathrm{d}x}\mathrm{d}\omega = N^{-1}\sum_{i=1}^{n} c_i \cdot \sup A_i = \\ = \sum_{i=1}^{n} m(A_i)\cdot \sup A_i. \tag{2.4}$$

Если случайная величина X является дискретной, то выражения для кумулятивных функций распределения и математических ожиданий остаются абсолютно такими же.

Пример 2.5. Предположим, что имеются шесть экспертных оценок ($N = 6$) о возможных значениях случайной величины X, характеризующей стоимость акций некоторой фирмы в условных единицах на следующий день, которая может меняться в интервале

[12] p-boxes (англ.)

[13] Подробный анализ p-блоков и алгоритмы работы с ними можно найти в разделе 9.7.

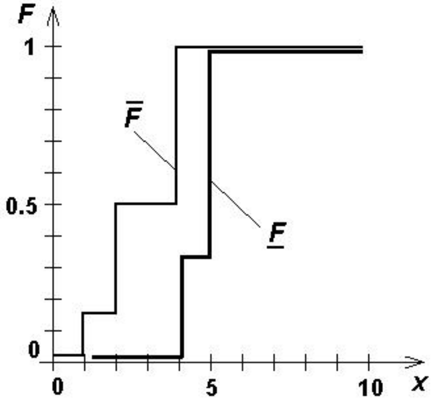

Рис. 2.3. Нижняя и верхняя функции распределения

$\Omega = [0, 10]$. Три эксперта ($c_1 = 3$) оценили интервал стоимости акций $A_1 = [4, 5]$, два эксперта ($c_2 = 2$) предоставили интервал $A_2 = [2, 4]$, и один эксперт ($c_3 = 1$) предоставил интервал $A_3 = [1, 5]$. Отсюда

$$m(A_1) = 1/2, \quad m(A_2) = 1/3, \quad m(A_3) = 1/6.$$

Графики нижней и верхней функций распределения показаны на рис. 2.3. Нижняя граница ожидаемой цены акций равна

$$\underline{\mathbb{E}}X = \frac{1}{2} \cdot 4 + \frac{1}{3} \cdot 2 + \frac{1}{6} \cdot 1 = 2.833,$$

а верхняя граница ожидаемой цены акций равна

$$\overline{\mathbb{E}}X = \frac{1}{2} \cdot 5 + \frac{1}{3} \cdot 4 + \frac{1}{6} \cdot 5 = 4.667.$$

Если рассматривать функцию $h(X)$ случайной величины X, то нижняя и верхняя границы математического ожидания функции h определяются на основе следующих выражений, обобщающих (2.3) и (2.4):

$$\underline{\mathbb{E}}h(X) = \sum_{i=1}^{n} m(A_i) \cdot \inf_{x \in A_i} h(x),$$

$$\overline{\mathbb{E}}h(X) = \sum_{i=1}^{n} m(A_i) \cdot \sup_{x \in A_i} h(x).$$

2.5. Основные отличия теории вероятностей и теории случайных множеств

Отметим основные отличия классической теории вероятностей и теории случайных множеств с точки зрения таких определяющих их понятий, как функция распределения и базовые вероятности.

Основное отличие между распределениями вероятностей и базовыми вероятностями заключается в том, что первые определены на множестве значений Ω, а вторые определены на множестве всех подмножеств множества Ω. Как следствие, базовые вероятности, функции доверия и правдоподобия имеют следующие свойства, отличающие их от элементов теории вероятностей.

1. В теории случайных множеств не требуется выполнения условия $m(\Omega) = 1$. В теории вероятностей условие $\Pr(\Omega) = 1$ должно выполняться обязательно. Более того, если $m(\Omega) = 1$, то семейство всех фокальных элементов есть $\mathcal{F} = \{\Omega\}$, т.е. имеется единственный фокальный элемент и этим элементом является само множество Ω. Это случай полного отсутствия какой-либо информации, и все возможные распределения вероятностей могут быть приписаны элементам множества Ω. В рамках же теории вероятностей, согласно принципу максимума энтропии, для описания отсутствия какой-либо информации используется равномерное распределение на Ω. Если имеется единственный фокальный элемент $A \subset \Omega$, то $m(A) = 1$ и $m(\Omega) = 0$.

2. В теории случайных множеств не требуется выполнения условия $m(A) \leq m(B)$, если $A \subseteq B$, и наоборот. В теории вероятностей данное условие должно выполняться обязательно. Более того, в теории случайных множеств каждый фокальный элемент должен рассматриваться как объект сам по себе. Если выполняется неравенство $m(A) \leq m(B)$, то это лишь означает, что объект A имеет меньшую вероятность, чем объект B, но при этом совсем не обязательно выполнение условия $A \subseteq B$.

3. В теории случайных множеств нет слишком жесткой связи между базовыми вероятностями $m(A)$ и $m(A^c)$, где A^c

– дополнение к A. В теории вероятностей всегда имеет место равенство $\Pr(A) = 1 - \Pr(A^c)$.

В теории случайных множеств аналогами основных свойств теории вероятностей являются следующие:

1. Пусть $\{A_i, i = 1, ..., n\}$ – множество всех фокальных элементов. Тогда выполняется условие $\sum_{i=1}^{n} m(A_i) = 1$.

2. Для любого события A выполняются условия $\mathrm{Bel}(A) = 1 - \mathrm{Pl}(A^c)$ и $\mathrm{Pl}(A) = 1 - \mathrm{Bel}(A^c)$.

2.6. Правила комбинирования свидетельств

Информация в виде различных наборов фокальных элементов с их базовыми вероятностями может быть получена из различных источников. Эти источники предоставляют различные данные об одном и том же объекте или явлении, но являются независимыми. Для комбинирования данных, полученных из независимых источников, используется ряд правил. При этом каждое правило имеет свои преимущества и недостатки. Достаточно детальный и критический обзор существующих правил комбинирования можно найти в работах [13, 74]. Поэтому рассмотрим только наиболее распространенные из них.

Предположим, что имеются два источника данных. Первый источник предоставляет N_1 наблюдений (свидетельств) $A_i^{(1)} \subseteq \Omega$, $i = 1, ..., n_1$, и $c_i^{(1)}$ – число наблюдаемых множеств $A_i^{(1)}$, $i = 1, ..., n_1$. Второй источник предоставляет N_2 наблюдений (свидетельств) $A_i^{(2)} \subseteq \Omega$, $i = 1, ..., n_2$, и $c_i^{(2)}$ – число наблюдаемых множеств $A_i^{(2)}$, $i = 1, ..., n_2$.

2.6.1. Правило комбинирования Демпстера

Правило комбинирования Демпстера основано на предположении, что источники данных абсолютно независимы. Обозначим базовые вероятности фокальных элементов, полученных из первого и второго источников, следующим образом:

$$m_1(A_i^{(1)}) = c_i^{(1)}/N_1, \; m_2(A_j^{(2)}) = c_j^{(2)}/N_2.$$

Тогда комбинированная базовая вероятность (m_{12}) вычисляется по формуле

$$m_{12}(A) = \frac{1}{1-K} \sum_{A_i^{(1)} \cap A_j^{(2)} = A} m_1(A_i^{(1)}) m_2(A_j^{(2)}),$$

где

$$K = \sum_{A_i^{(1)} \cap A_j^{(2)} = \varnothing} m_1(A_i^{(1)}) m_2(A_j^{(2)})$$

и $m_{12}(\varnothing) = 0$. Заметим, что если два фокальных элемента из разных источников не пересекаются, т.е. $A_i^{(1)} \cap A_j^{(2)} = \varnothing$, то соответствующие источники противоречат друг другу или конфликтуют. Например, если один эксперт считает, что температура воздуха на следующий день будет в интервале от 10 до 12 градусов, а другой эксперт полагает, что температура будет в пределах от 16 до 20 градусов, то становится очевидным, что экспертные оценки противоречивы или конфликтны, так как $[10, 12] \cap [16, 20] = \varnothing$. Конфликтность свидетельств учитывается коэффициентом K, который представляет собой общую базовую вероятность, связанную с конфликтными свидетельствами. Если все свидетельства противоречивы, т.е. $K = 1$, то полностью противоречивые источники не могут быть объединены при помощи правила комбинирования Демпстера.

Пример 2.6. Четыре предприятия ($\Omega = \{1, 2, 3, 4\}$) являются кандидатами для покупки акций. 5 экспертов из первой группы считают, что необходимо покупать акции первого предприятия; 3 эксперта из первой группы считают, что необходимо покупать акции второго или третьего предприятия; 8 экспертов из второй независимой группы считают, что необходимо покупать акции первого или второго предприятия; 7 экспертов из второй группы считают, что необходимо покупать акции третьего предприятия, и один эксперт из второй группы считает, что необходимо покупать акции четвертого предприятия. Первый источник: $N_1 = 8$, $c_1^{(1)} = 5$, $A_1^{(1)} = \{1\}$, $m_1(A_1^{(1)}) = 5/8$, $c_2^{(1)} = 3$, $A_2^{(1)} = \{2, 3\}$, $m_1(A_2^{(1)}) = 3/8$. Второй источник: $N_2 = 16$, $c_1^{(2)} = 8$, $A_1^{(2)} = \{1, 2\}$, $m_2(A_1^{(2)}) = 8/16$, $c_2^{(2)} = 3$, $A_2^{(2)} = \{3\}$, $m_2(A_2^{(2)}) = 7/16$, $c_3^{(2)} = 1$, $A_3^{(2)} = \{4\}$, $m_2(A_3^{(2)}) = 1/16$. В табл. 2.1 представлены все возможные пересечения $A_i^{(1)} \cap A_j^{(2)}$ фокальных элементов из двух источников.

Таблица 2.1. Данные, полученные из двух источников, и их пересечения

		$A_i^{(1)}$	
		$\{1\}$	$\{2,3\}$
$A_i^{(2)}$	$\{1,2\}$	$\{1\}$	$\{2\}$
	$\{3\}$	$\varnothing$	$\{3\}$
	$\{4\}$	$\varnothing$	$\varnothing$

Коэффициент конфликтности K вычисляется, исходя из того что $A_1^{(1)} \cap A_2^{(2)} = \varnothing$, $A_1^{(1)} \cap A_3^{(2)} = \varnothing$, $A_2^{(1)} \cap A_3^{(2)} = \varnothing$ (см. в табл. 2.1 клетки с элементами "$\varnothing$"), т.е.

$$K = m_1(A_1^{(1)})m_2(A_2^{(2)}) + m_1(A_1^{(1)})m_2(A_3^{(2)}) + \\ + m_1(A_2^{(1)})m_2(A_3^{(2)}) = \\ = \frac{5}{8} \cdot \frac{7}{16} + \frac{5}{8} \cdot \frac{1}{16} + \frac{3}{8} \cdot \frac{1}{16} = 0.336$$

Отсюда $1 - K = 1 - 0.336 = 0.664$. Из табл. 2.1 также видно, что непустые пересечения имеют вид $\{1\}$, $\{2\}$, $\{3\}$. Тогда

$$m_{12}(\{1\}) = m_1(A_1^{(1)})m_2(A_1^{(2)})/0.664 = 0.4706,$$

$$m_{12}(\{2\}) = m_1(A_2^{(1)})m_2(A_1^{(2)})/0.664 = 0.2824,$$

$$m_{12}(\{3\}) = m_1(A_2^{(1)})m_2(A_2^{(2)})/0.664 = 0.2470.$$

Из полученных базовых вероятностей можно вычислить функции доверия и правдоподобия:

$$\mathrm{Bel}(\{1\}) = m_{12}(\{1\}) = \mathrm{Pl}(\{1\}) = 0.4706,$$

$$\mathrm{Bel}(\{2\}) = m_{12}(\{2\}) = \mathrm{Pl}(\{2\}) = 0.2824,$$

$$\mathrm{Bel}(\{3\}) = m_{12}(\{3\}) = \mathrm{Pl}(\{3\}) = 0.2470.$$

В отзыве на книгу Шейфера "Математическая теория свидетельств" Заде представил пример, показывающий, что правило комбинирования Демпстера может давать "некорректные" результаты в случае большого количества противоречивых данных. Представим, что пациент осматривается двумя врачами-терапевтами по поводу неврологических симптомов. Первый врач уверен, что у пациента либо менингит с вероятностью 0.99, либо опухоль мозга с вероятностью 0.01. Второй врач уверен, что у пациента расстройство нервной системы после сотрясения с вероятностью 0.99, но допускает опухоль мозга с вероятностью 0.01. Комбинируя полученные свидетельства при помощи правила Демпстера, получаем равенство

$$\mathrm{Bel}(\text{опухоль}) = m(\text{опухоль}) = 1.$$

Следовательно, результат комбинирования поддерживает полностью диагноз, который оба врача рассматривают как очень маловероятный. Необходимо отметить, что причиной такого результата является предположение о том, что источники абсолютно надежны и предоставляют верные сведения. Поэтому при наличии противоречивых оценок из различных источников правило Демпстера "ищет" то общее, что имеется в обоих источниках, и "отбрасывает" все, что различается в них.

2.6.2. Правило дисконтирования

Как было отмечено выше, правило комбинирования Демпстера предполагает абсолютную надежность источников информации. Однако существует всегда сомнение, что источники абсолютно надежны. Для того чтобы учесть надежность источников, Шейфер предложил использование дисконтирования базовых вероятностей некоторым коэффициентом $\alpha \in [0,1]$, характеризующим надежность источника, т.е. умножения базовой вероятности на α. В результате для каждого фокального элемента A получаем новые базовые вероятности $m^{\alpha}(A) = (1-\alpha)m(A)$. Если коэффициент дисконтирования равен 1, то источник является абсолютно ненадежным и $m^{\alpha}(A) = 0$. Наоборот, если $\alpha = 0$, то источник является абсолютно надежным и $m^{\alpha}(A) = m(A)$. Чтобы выполнить условие нормирования базовых вероятностей (сумма базовых вероятностей всех фокальных элементов равна 1), при дисконтировании добавляется базовая вероятность всего множества Ω, т.е. $m^{\alpha}(\Omega) = \alpha + (1-\alpha)m(\Omega)$. Фактически добавление ненулевой базовой вероятности Ω не меняет информации, имеющейся в распоряжении. Если эксперт говорит, что "любой элемент

Ω" может быть "истинным" значением случайной величины, то он не дает никакой дополнительной информации.

Интересно отметить, что использование дисконтирования даже при очень малых значениях α делает коэффициент K в правиле комбинирования Демпстера не равным 1 и, следовательно, позволяет всегда найти комбинированную оценку независимо от количества противоречивой информации[14].

Пример 2.7. Вернемся к предыдущему примеру о покупке акций четырех предприятий. 100 экспертов из первой группы считают, что необходимо покупать акции второго предприятия; 60 экспертов из первой группы считают, что необходимо покупать акции второго или третьего предприятия; 8 экспертов из второй независимой группы считают, что необходимо покупать акции первого. Первый источник: $N_1 = 160$, $c_1^{(1)} = 100$, $A_1^{(1)} = \{2\}$, $m_1(A_1^{(1)}) = 100/160$, $c_2^{(1)} = 60$, $A_2^{(1)} = \{2,3\}$, $m_1(A_2^{(1)}) = 60/160$. Второй источник: $N_2 = 8$, $c_1^{(2)} = 8$, $A_1^{(2)} = \{1\}$, $m_2(A_1^{(2)}) = 1$. Нетрудно увидеть, что

$$K = m_1(A_1^{(1)})m_2(A_1^{(2)}) + m_1(A_2^{(1)})m_2(A_1^{(2)}) = 1.$$

Следовательно, невозможно получить комбинированную оценку, используя правило комбинирования Демпстера. Поэтому используем правило дисконтирования с учетом надежности источников. Так как первый источник содержит намного больше экспертов, чем второй, то можно считать его более надежным по сравнению с первым источником[15]. Примем $\alpha_1 = 1 - 160/168 = 0.048$ и $\alpha_2 = 1 - 8/168 = 0.952$. Заметим, что вовсе не обязательно, чтобы сумма коэффициентов дисконтирования была бы равна 1. Тогда

$$m_1^{\alpha_1}(A_1^{(1)}) = (1-\alpha_1)\cdot 100/160 = 0.595,$$

$$m_1^{\alpha_1}(A_2^{(1)}) = (1-\alpha_1)\cdot 60/160 = 0.357,$$

$$m_1^{\alpha_1}(\Omega) = \alpha_1 = 0.048,$$

[14]Одной из причин этого факта является ненулевая базовая вероятность Ω. С одной стороны, оценка Ω не дает никакой дополнительной информации. С другой стороны, она размывает конечный результат и за счет этого делает его более осторожным.

[15]Подход для анализа коэффициентов дисконтирования, используемый в примере, основан на подсчете относительного соотношения количества оценок каждого источника. Этот подход не является строгим с математической точки зрения, но может применяться в тех случаях, когда нет никакой дополнительной информации о надежности источников.

$$m_2^{\alpha_2}(A_1^{(2)}) = (1-\alpha_2)\cdot 1 = 0.048,$$
$$m_2^{\alpha_2}(\Omega) = \alpha_2 = 0.952.$$

Теперь можно использовать правило комбинирования Демпстера, согласно которому

$$K = m_1^{\alpha_1}(A_1^{(1)})m_2^{\alpha_2}(A_1^{(2)}) + m_1^{\alpha_1}(A_2^{(1)})m_2^{\alpha_2}(A_1^{(2)}) = \\ = 0.595\cdot 0.048 + 0.357\cdot 0.048 = 0.046,$$

$$1-K = 0.954,$$

$$m_{12}(\{1\}) = m_1^{\alpha_1}(\Omega)m_2^{\alpha_2}(A_1^{(2)})/(1-K) = \\ = 0.048\cdot 0.048/0.954 = 0.002,$$

$$m_{12}(\{2\}) = m_1^{\alpha_1}(A_1^{(1)})m_2^{\alpha_2}(\Omega)/(1-K) = \\ = 0.595\cdot 0.952/0.954 = 0.594,$$

$$m_{12}(\{2,3\}) = m_1^{\alpha_1}(A_2^{(1)})m_2^{\alpha_2}(\Omega)/(1-K) = \\ = 0.357\cdot 0.952/0.954 = 0.356,$$

$$m_{12}(\Omega) = m_1^{\alpha_1}(\Omega)m_2^{\alpha_2}(\Omega)/(1-K) = \\ = 0.048\cdot 0.952/0.954 = 0.048.$$

Следует отметить, что

$$m_{12}(\{1\}) + m_{12}(\{2\}) + m_{12}(\{2,3\}) + m_{12}(\Omega) = 1.$$

Найдем функции доверия и правдоподобия для всех предприятий

$$\mathrm{Bel}(\{1\}) = m_{12}(\{1\}) = 0.002,$$
$$\mathrm{Pl}(\{1\}) = m_{12}(\{1\}) + m_{12}(\Omega) = 0.05,$$

$$\mathrm{Bel}(\{2\}) = m_{12}(\{2\}) = 0.594,$$
$$\mathrm{Pl}(\{2\}) = m_{12}(\{2\}) + m_{12}(\{2,3\}) + m_{12}(\Omega) = 0.998,$$

$$\mathrm{Bel}(\{3\}) = 0,\ \mathrm{Pl}(\{3\}) = m_{12}(\{2,3\}) + m_{12}(\Omega) = 0.404,$$

$$\mathrm{Bel}(\{4\}) = 0,\ \mathrm{Pl}(\{4\}) = m_{12}(\Omega) = 0.048.$$

2.6.3. Правило комбинирования Ягера

Одной из альтернатив рассмотренным выше правилам является правило предложенное Ягером [144]. Он отметил, что важным положительным свойством ряда правил комбинирования свидетельств является их способность модифицировать уже комбинированную структуру, когда появляется новая информация. Это свойство также называется ассоциативностью правил. В то же время во многих случаях именно отсутствие этого свойства имеет положительные стороны. Например, арифметическое среднее, как правило комбинирования, не является ассоциативным в том смысле, что, имея в распоряжении некоторое среднее, нельзя его модифицировать при получении новой статистической информации. Однако арифметическое среднее можно модифицировать, добавляя новые точечные данные к сумме уже существующих данных и разделив новую сумму на общее число данных. Это свойство называется квазиассоциативностью, на основе которого Ягер разработал правило комбинирования свидетельств.

Ягер ввел так называемую комбинированную “универсальную” вероятность q, определяемую как

$$q(A) = \sum_{B \cap C = A} m_1(B) m_2(C),$$

где A – пересечение подмножеств $B \in \mathcal{P}o(\Omega)$ и $C \in \mathcal{P}o(\Omega)$.

Комбинированная вероятность q может использоваться для любого числа свидетельств. Пусть $m_1, ..., m_n$ – базовые вероятности n источников данных и F_i – множество фокальных элементов, соответствующих i-му источнику данных, а $A_k^{(i)}$ – один из элементов F_i. Тогда правило комбинирования определяется как

$$q(A) = \sum_{A_i^{(1)} \cap A_j^{(2)} \cap \cdots \cap A_k^{(n)} = A} m_1(A_i^{(1)}) m_2(A_j^{(2)}) \cdots m_n(A_k^{(n)}).$$

Одним из очевидных отличий правила комбинирования Ягера от ранее рассмотренных правил является отсутствие коэффициента нормализации $1 - K$. В правиле комбинирования Ягера его удалось избежать благодаря тому, что комбинированная вероятность q пустого множества может быть больше 0, т.е. $q(\varnothing) \geq 0$. При этом $q(\varnothing)$ вычисляется абсолютно так же как коэффициент конфликтности источников K в правиле комбинирования Демпстера. Затем Ягер добавляет значение $q(\varnothing)$ к комбинированной "универсальной" вероятности $q(\Omega)$ всего множества Ω для получения комбинированной базовой вероятности

$$m_{\text{Yag}}(\Omega) = q(\Omega) + q(\varnothing).$$

Следовательно, вместо нормализации с учетом противоречивости источников, как это было сделано в правиле комбинирования Демпстера, Ягер переносит противоречивость, или конфликтность, свидетельств на все множество Ω. При этом базовая вероятность $m_{\text{Yag}}(\Omega)$ может теперь интерпретироваться как степень незнания или неопределенности.

Базовые вероятности для других множеств определяются как

$$m_{\text{Yag}}(\varnothing) = 0, \quad m_{\text{Yag}}(A) = q(A), \; A \neq \varnothing, \; A \neq \Omega.$$

Базовые вероятности, полученные с использованием правила комбинирования Демпстера, могут быть выражены через комбинированные "универсальные" вероятности q, полученные на основе правила комбинирования Ягера, следующим образом:

$$\begin{aligned} &m(\varnothing) = 0, \quad m(\Omega) = q(\Omega)/(1 - q(\varnothing)), \\ &m(A) = q(A)/(1 - q(\varnothing)), \end{aligned}$$

где $A \neq \varnothing$, $A \neq \Omega$.

Пример 2.8. Вернемся снова к примеру о покупке акций четырех предприятий. Так как $q(\varnothing)$ вычисляется абсолютно так же как коэффициент противоречивости, или конфликтности, K в правиле Демпстера, то $q(\varnothing) = 0.336$. Отсюда

$$\begin{aligned} m_{\text{Yag}}(\{1\}) = q(\{1\}) = m_1(A_1^{(1)})m_2(A_1^{(2)}) = \\ = \frac{5}{8} \cdot \frac{8}{16} = 0.3125, \end{aligned}$$

$$m_{\mathrm{Yag}}(\{2\}) = q(\{2\}) = m_1(A_2^{(1)})m_2(A_1^{(2)}) = \\ = \frac{3}{8}\cdot\frac{8}{16} = 0.1875,$$

$$m_{\mathrm{Yag}}(\{3\}) = q(\{3\}) = m_1(A_2^{(1)})m_2(A_2^{(2)}) = \\ = \frac{3}{8}\cdot\frac{7}{16} = 0.1641,$$

$$q(\Omega) = 0,\ m_{\mathrm{Yag}}(\Omega) = q(\Omega) + q(\varnothing) = 0.336.$$

Из полученных базовых вероятностей можно вычислить функции доверия и правдоподобия:

$$\mathrm{Bel}(\{1\}) = m_{\mathrm{Yag}}(\{1\}) = 0.3125,$$
$$\mathrm{Pl}(\{1\}) = m_{\mathrm{Yag}}(\{1\}) + m_{\mathrm{Yag}}(\Omega) = 0.6485,$$

$$\mathrm{Bel}(\{2\}) = m_{\mathrm{Yag}}(\{2\}) = 0.1875,$$
$$\mathrm{Pl}(\{2\}) = m_{\mathrm{Yag}}(\{2\}) + m_{\mathrm{Yag}}(\Omega) = 0.5235,$$

$$\mathrm{Bel}(\{3\}) = m_{\mathrm{Yag}}(\{3\}) = 0.1641,$$
$$\mathrm{Pl}(\{3\}) = m_{\mathrm{Yag}}(\{3\}) + m_{\mathrm{Yag}}(\Omega) = 0.5.$$

2.6.4. Параметрическое семейство правил комбинирования Инагаки

Используя особенности правил комбинирования Демпстера и Ягера, Инагаки [49] предложил параметрическое семейство правил. При этом он ввел такую функцию $f(A)$, что для непустого множества A можно записать комбинированную базовую вероятность в виде

$$m(A) = q(A) + f(A)\cdot q(\varnothing)$$

и

$$\sum_{A\subset\Omega,\ A\neq\varnothing} f(A) = 1, \quad f(A) \geq 0.$$

Функцию f можно интерпретировать как функцию масштаба для $q(\varnothing)$, и степень конфликта определяется параметром $k = f(A)/q(A)$. Инагаки ограничил класс правил комбинирования условием

$$\frac{m(A)}{m(C)} = \frac{q(A)}{q(C)}, \quad A, C \neq \varnothing,\ A, C \neq \Omega.$$

Сохранение отношения между m и q означает, что нет каких-либо "метасведений" о надежности или качестве источников информации или экспертов. Если применить некоторый весовой множитель к свидетельствам на основе какой-либо дополнительной информации об источниках, то это изменит приведенное отношение и равенство не будет выполняться. Поэтому правило Инагаки может быть применено только в случае отсутствия какой-либо информации об источниках.

Если теперь обозначить комбинированную базовую вероятность m_{In}, то преобразование приведенных выше выражений дает следующие результаты:

$$m_{\mathrm{In}}(A) = [1 + kq(\varnothing)]q(A),\ A \neq \varnothing,\ A \neq \Omega,$$

$$m_{\mathrm{In}}(\Omega) = [1 + kq(\varnothing)]q(\Omega) + [1 + kq(\varnothing) - k]q(\varnothing),$$

$$0 \leq k \leq \frac{1}{1 - q(\varnothing) - q(\Omega)}.$$

Параметр k используется для нормализации, и его задание является важным этапом при реализации правила комбинирования, хотя обоснованные подходы для выбора k отсутствуют в литературе. Если $k = 0$, то правило комбинирования Инагаки совпадает с правилом Ягера. Если $k = 1/(1 - q(\varnothing))$, то получаем правило комбинирования Демпстера. Таким образом, правило комбинирования Инагаки можно рассматривать как параметрическое семейство правил с параметром k.

2.7. Заключение

Теория Демпстера–Шейфера является достаточно мощным инструментом для моделирования неточности и неопределенности. Она позволяет моделировать полное отсутствие исходной информации. Функции доверия и правдоподобия некоторого события можно рассматривать как верхнюю и нижнюю границы вероятности этого события. С этой точки зрения теория Демпстера–Шейфера не является альтернативой теории вероятностей, а дополняет и обобщает ее.

Несмотря на ряд достоинств рассмотренной теории, она моделирует только определенный вид исходной информации.

Кроме того, определение базовых вероятностей с использованием правила (2.1) может привести к противоречивым результатам в задачах принятия решений. Одной из причин этого является тот факт, что возможное отсутствие априорной информации (наблюдения, оценки) для части элементов универсального множества Ω делает функции доверия и правдоподобия этих элементов нулевыми. Это говорит о том, что появление таких элементов абсолютно невозможно, что в большинстве практических задач является неверным и обусловлено ограниченностью выборки. Если при большом числе наблюдений определенная часть элементов множества Ω никак бы не проявлялась, то можно говорить о нулевой вероятности появления этих элементов. Однако для ограниченной выборки данное утверждение является слишком рискованным. Интересно отметить, что единичный коэффициент конфликтности в правиле комбинирования Демпстера также является следствием отмеченного недостатка. Поэтому для устранения отмеченных недостатков требуются новые методы описания неопределенности и неточности, а также модификации теория Демпстера–Шейфера, которые будут рассмотрены ниже.