مروری بر مبانی نظریه دمپستر– شفر

بابک نجار اعرابی [۱]، طاهر شهبازی میرزا حسنلو[۱]

araabi@ut.ac.ir

چکیده

این مقاله به مرور مباحث، تعاریف، قضایا و نتایج اولیه و اساسی در نظریه دمپستر–شفر یا همان نظریه شواهد می‌پردازد. ابتدا نظریه حدود پایین و بالا احتمال آنچنان که توسط دمپستر ارائه شده است معرفی می‌گردد. سپس نظریه شواهد شفر در یک ساختار اصل موضوعی معرفی شده، نشان می‌دهیم که نتایج آن معادل با نظریه حدود پایین و بالا احتمال دمپستر است. در این راستا توابع باور و موجه‌نمایی متناظر با حدود پایین و بالای احتمال معرفی خواهند شد. به علاوه به معرفی مجموعه توزیع‌های احتمال سازگار با یک ساختار اعتقادی می‌پردازیم و نشان می‌دهیم که حدود بالا و پایین احتمال هر پیشامد را می‌توان با بهینه‌سازی احتمال پیشامد روی فضای شامل مجموعه‌ای از توزیع‌های سازگار با ساختار اعتقادی به‌دست آورد. بحثی درباره اعمال شرط و نیز ترکیب اطلاعات حاصل از منابع مستقل در قالب نظریه شواهد خواهیم کرد و با اشاره‌ای مختصر به آنچه در باره‌ی نظریه دمپستر–شفر مهم است اما در این مقاله فرصت پرداختن به آن نبوده، بحث را خاتمه می‌دهیم. در جای جای مقاله با ذکر مثال‌های کاربردی سعی در تشریح مباحث نظری شده است.

واژه‌های کلیدی: نظریه دمپستر–شفر، نظریه شواهد، مجموعه توزیع‌های سازگار با یک ساختار اعتقادی، تابع باور، تابع موجه‌نمایی، حدود پایین و بالای احتمال، عدم آگاهی، قاعده ترکیب دمپستر، قاعده اعمال شرط دمپستر.

۱ مقدمه

اواخر دهه ۶۰ میلادی، دمپستر[۲] در مقاله مشهور خود [۱] دیدگاه جدیدی را در بیان اندازه احتمال روی یک فضا چون Ω ارائه داد. او فرض کرد که اندازه احتمال μ روی فضای X موجود باشد و به‌علاوه اعضای X به‌وسیله یک نگاشت چند مقداری[۳] Γ به روی زیرمجموعه‌های Ω تصویر شوند، یعنی

$$\Gamma : X \longrightarrow P_\Omega \quad x \longrightarrow \Gamma(x) \qquad (۱)$$

که در آن P_Ω مجموعه تمام زیرمجموعه‌های Ω، و $\Gamma(x)$ نگاشت $x \in X$ تحت Γ است (شکل ۱).

[۱] دانشکده مهندسی برق و کامپیوتر، دانشگاه تهران
[۲] A.P. Dempster
[۳] Multi-valued mapping

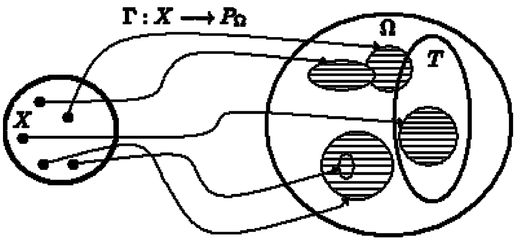
حال سوال این است چه قضاوتی در مورد احتمال یک پیشامد T روی Ω می‌توان کرد؟ در این مقاله $\Omega = \{w_۱, w_۲, \ldots, w_n\}$ مجموعه‌ای با تعداد اعضای متناهی فرض شده است، لذا ملاحظات اندازه‌پذیری موضوعیت ندارند.

اگر کمی متفاوت به نگاشت Γ نگاه کنیم درک ماهیت آن ساده‌تر خواهد بود. نظیر Γ می‌توانیم یک رابطه γ روی $X \times \Omega$ تعریف کنیم آن‌چنان‌که

$$\forall (x,w) \in X \times \Omega \, : \, x\gamma w \Longleftrightarrow \omega \in \Gamma(x) \qquad (۲)$$

در حالت خاصی که رابطه γ یک رابطه تابعی یک به یک باشد محاسبه احتمال پیشامد $T \subset \Omega$ ساده است و با محاسبه وارون

تابع انجام می‌شود. اما وقتی در حالت کلی γ تنها یک رابطه روی $X \times \Omega$ باشد محاسبه احتمال پیشامد T درگیر با نوعی نایقینی[۴] می‌شود، چراکه مطمئن نیستیم چگونه احتمال $x \in X$ را، که تحت نگاشت Γ به $\Gamma(x)$ منتقل شده، بین اعضای مجموعه $\Gamma(x)$ تقسیم کنیم. رابطه γ را معمولاً یک رابطه سازگاری[۵] بین اعضاء X و Ω می‌نامند و $x\gamma w$ به صورت «x با w سازگار است» تعبیر می‌شود. مثال ۱ تعابیر فوق را روشن‌تر می‌کند.



شکل ۱. نگاشت چندمقداری از X به Ω و پیشامد $T \subset \Omega$

مثال ۱ یک موسسه مالی و اعتباری قصد سرمایه‌گذاری در حوزه تولید محصولات غذایی را دارد. این کار از طریق پرداخت وام‌های بلند مدت به شرکت‌های متقاضی انجام خواهد شد. تعداد زیادی شرکت متقاضی دریافت وام هستند و موسسه مالی برای تصمیم‌گیری نیاز دارد که بداند آیا هر شرکت متقاضی در حال حاضر در وضعیت نزدیک به ورشکستگی (w_1)، سوددهی ناپایدار (w_2)، و یا سوددهی پایدار (w_3) است. شرکت‌های متقاضی اما معمولاً این اطلاعات را از اسرار خود می‌دانند و در اختیار قرار نمی‌دهند. موسسه مالی به یک موسسه حسابرسی که در حوزه شرکت‌های تولیدی محصولات غذایی مطالعه می‌کند توسل می‌جوید. مطالعات موسسه حسابرسی نشان می‌دهد که ارتباطی بین وضعیت سوددهی شرکت‌ها و میزان تولید ۶ ماه آینده آن‌ها وجود دارد. میزان تولید نسبت به میزان تولید اسمی در پنج دسته طبقه‌بندی می‌شود: کمتر از ۲۰ درصد (x_1)، بین ۲۰ تا ۵۰ درصد (x_2)، بین ۵۰ تا ۸۰ درصد (x_3)، بین ۸۰ تا ۱۰۰ درصد (x_4) و بیش از ۱۰۰ درصد (x_5) تولید اسمی.

جدول ۱ رابطه وضعیت سوددهی و میزان تولید را به تشخیص موسسه حسابرسی نشان می‌دهد.

جدول ۱. رابطه (γ) وضعیت سوددهی و میزان تولید (۱ به معنی وجود رابطه و ۰ به معنی عدم وجود آن است)

رابطه γ	x_1	x_2	x_3	x_4	x_5
w_1	۱	۱	۰	۰	۱
w_2	۰	۱	۰	۱	۱
w_3	۰	۰	۱	۱	۰

جدول ۲. اندازه احتمال روی فضای X

توزیع μ	x_1	x_2	x_3	x_4	x_5
$\mu(x_i)$	۰/۷	۰/۲	۰/۱	۰	۰

حال اگر برای شرکتی میزان تولید ۶ ماه آینده با توزیع جدول ۲ برآورد شده باشد، آن گاه به نظر می‌رسد بتوان با داشتن γ و μ به محاسبه احتمال‌های w_1، w_2 و w_3 پرداخت. $\mu(x_1)$ به w_1 منتقل می‌شود و $\mu(x_3)$ به w_3؛ اما $\mu(x_2)$ می‌تواند به هر نسبتی به هر دوی w_1 و w_2 منتقل شود چرا که با توجه به رابطه γ می‌دانیم x_2 با هر دوی w_1 و w_2 سازگار است. دیده می‌شود که محاسبه احتمال w_1 ها با نوعی نایقینی درگیر است. دمپستر نشان داد که تعبیر احتمال یک پیشامد روی Ω درگیر با نوعی عدم آگاهی[۶] است؛ به این ترتیب که به جای آنکه به یک عدد به‌عنوان احتمال پیشامد برسیم، بازه‌ای از اعداد —محدود به حدود پایین و بالا— وجود دارد که می‌توانند احتمال یک پیشامد باشند. به این ترتیب دو اندازه μ_l و μ_h روی Ω به‌دست می‌آید که برای آنها، به ترتیب، تعبیر حدود

Uncertainty[۴]
Compatibility relation[۵]
Ignorance[۶]

پایین و بالا احتمال [۷] را به‌کار می‌بریم. اما این اندازه‌ها، اندازه احتمال نیستند. نه خاصیت جمع‌پذیری احتمال در مورد آن‌ها برقرار است و نه خاصیت بهنجار بودن، و داریم $(T, R \subset \Omega)$

$$\mu_l(T) + \mu_l(R) \leq \mu_l(T \cup R) + \mu_l(T \cap R) \qquad (۳)$$

$$\mu_l(T) + \mu_l(T^c) \leq ۱ \qquad (۴)$$

$$\mu_h(T) + \mu_h(R) \geq \mu_h(T \cup R) + \mu_h(T \cap R) \qquad (۵)$$

$$\mu_h(T) + \mu_h(T^c) \geq ۱ \qquad (۶)$$

اما μ_l و μ_h هر دو دارای خاصیت یکنوایی هستند

$$\begin{matrix} \mu_l(T) \leq \mu_l(R) \\ \mu_h(T) \leq \mu_h(R) \end{matrix} \quad \forall T, R: \ T \subset R \subset \Omega \qquad (۷)$$

و نیز

$$\mu_l(\emptyset) = \mu_h(\emptyset) = ۰ \qquad \mu_l(\Omega) = \mu_h(\Omega) = ۱ \qquad (۸)$$

پس می‌توان μ_l و μ_h را در ساختار کلی‌تر اندازه‌های فازی [۸] قرار داد [۲] که اندازه احتمال هم حالت خاصی از آن‌ها است. اگر به این نکته توجه کنیم که دمپستر در مراحل شکل‌گیری این نظریه مباحثات زیادی با سویج [۹] داشته‌است [۳]، بیشتر به دغدغه‌هایی که منجر به ارائه نظریه شده‌اند پی می‌بریم. آنها درگیر این مسئله بودند که ساختار کلاسیک احتمال، امکان نمایش مناسب عدم آگاهی را به‌دست نمی‌دهد. در حالی که حداقل در بیان احتمال شخصی و ذهنی [۱۰] اکثرا با نوعی عدم آگاهی و نایقینی مواجه‌ایم. به علاوه احتمال یک اندازه جمع‌پذیر است، در حالی که مطالعات تجربی چنین خاصیتی را در مورد احتمال شخصی و ذهنی نشان نمی‌دهد و بیشتر به نظر می‌آید احتمال ذهنی یک اندازه ترتیبی است. نظریه دمپستر پاسخی برای این دو مسئله به‌دست می‌دهد. دو اندازه ترتیبی و غیر جمع‌پذیر ارائه می‌شوند که بیانگر محدوده احتمال هستند، همراه با مدلی که امکان حمل عدم آگاهی را نیز دارد.

بعدها در نیمه دوم دهه ۷۰ میلادی، شفر [۱۱] نظریه شواهد خود را ارائه داد که توسعه طبیعی نظریه احتمال بیزی است [۴]، و نتایج آن در انطباق کامل با نظریه دمپستر قرار دارد. مجموعه این نتایج را، با هر دو تعبیر ممکن نظریه دمپستر—شفر [۱۲] یا همان نظریه شواهد [۱۳] می‌نامند.

به علاوه دمپستر در همان مقاله [۱] از زاویه‌ای متفاوت به مسئله حدود پایین و بالا احتمال پرداخته و آنها را به‌عنوان حاصل فرایند کمینه و بیشینه‌سازی احتمال پیشامد روی مجموعه‌ای از توزیع‌های احتمال که به طور یکسان ممکن [۱۴] هستند، مطرح می‌کند. این دیدگاه گر چه کم مورد توجه قرار گرفته است اما می‌تواند نقطه شروع مناسبی برای توسعه روش‌های استنباط و تصمیم‌گیری آماری مبتنی بر نظریه دمپستر—شفر باشد.

در ادامه این مقاله، ابتدا در بخش ۲ نظریه حدود پایین و بالا احتمال را آن چنان که دمپستر معرفی کرد بیان می‌کنیم، و سپس در بخش ۳ دیدگاه شفر را ارائه می‌دهیم و خواهیم دید که به نتایج مشابهی می‌رسیم. در بخش ۴ قاعده اعمال شرط دمپستر [۱۵] و قاعده ترکیب دمپستر [۱۶] را معرفی می‌کنیم. در بخش ۵ به معرفی مجموعه توزیع‌های احتمال سازگار با یک ساختار

[۷] Upper and lower probabilities
[۸] Fuzzy measures
[۹] L.J. Savage
[۱۰] Personal and subjective probability
[۱۱] G. Shafer
[۱۲] Dempster-Shafer Theory
[۱۳] Evidence theory
[۱۴] Feasible
[۱۵] Dempster's rule of conditioning
[۱۶] Dempster's rule of combination

اعتقادی [۱۷] و بهینه‌سازی احتمال پیشامد روی آن می‌پردازیم، و سرانجام در بخش ٦ با اشاره‌ای به برخی دستاوردها و نتایج مهم در زمینه نظریه شواهد به جمع‌بندی نهایی می‌پردازیم.

۲ نظریه حدود پایین و بالا احتمال دمپستر

فضای X همراه با اندازه احتمال μ را در نظر می‌گیریم. μ به هر زیرمجموعه X چون A، عدد $\mu(A)$ را به عنوان احتمال پیشامد A نسبت می‌دهد. حال فضای Ω را در نظر می‌گیریم که از طریق نگاشت چند مقداری Γ با X مربوط است (شکل ۱). می‌خواهیم احتمال پیشامد $T \subset \Omega$ را بیابیم. به این منظور دو مجموعه $T_*, T^* \subset X$ را به صورت زیر تعریف می‌کنیم

$$T_* = \{x | x \in X, \Gamma x \neq \emptyset, \Gamma x \subset T\} \qquad (۹)$$

$$T^* = \{x | x \in X, \Gamma x \cap T \neq \emptyset\} \qquad (۱۰)$$

در حالت خاص $T = \Omega$ داریم

$$\Omega^* = \Omega_* = \{x | x \in X, \Gamma x \neq \emptyset\} \qquad (۱۱)$$

که همان دامنه نگاشت Γ است.

حال اندازه‌های μ_l و μ_h روی Ω به صورت زیر تعریف می‌شوند

$$\mu_l(T) = \frac{\mu(T_*)}{\mu(\Omega^*)} \qquad (۱۲)$$

$$\mu_h(T) = \frac{\mu(T^*)}{\mu(\Omega^*)} \qquad (۱۳)$$

تقسیم بر $\mu(\Omega^*)$ نقش بهنجار کردن اندازه را دارد و طبعا با فرض $\mu(\Omega^*) \neq 0$ امکان‌پذیر است. با توجه به تعریف T_* و T^* به سادگی می‌توان دید که

$$T_* \subset T^* \longrightarrow \mu_l(T) \leq \mu_h(T) \quad \forall T \subset \Omega \qquad (۱۴)$$

در واقع $\mu(T^*)$ تعبیر بیشترین مقدار احتمالی که اندازه احتمال μ تحت نگاشت Γ می‌تواند به T (یا بهتر بگوییم عضوهای T) منتقل کند را دارد و به همین ترتیب $\mu(T_*)$ تعبیر کمترین مقدار احتمالی که تحت Γ از μ به T منتقل می‌شود را دارد. پس بازه $[\mu_l(T), \mu_h(T)]$ تعبیر مجموعه مقادیر ممکن برای احتمال پیشامد T روی Ω را می‌یابد

$$P(T) \in [\mu_l(T), \mu_h(T)] \longrightarrow \mu_l(T) \leq P(T) \leq \mu_h(T) \forall T \subset \Omega \qquad (۱۵)$$

و طول بازه مقادیر ممکن احتمال، $\mu_h(T) - \mu_l(T)$، بیانگر میزان عدم آگاهی یا نایقینی در بیان احتمال پیشامد T روی Ω است. در ادامه با معرفی اعضای کانونی و جرم نظیر آنها محاسبه حدود بالا و پایین احتمال را بدون ارجاع مستقیم به فضای X و نگاشت Γ انجام خواهیم داد. نظیر هر عضو در برد Γ می‌توان مجموعه‌ای از اعضای X را یافت که بدان نظیر شود. اگر برد Γ دارای l عضو متمایز و ناتهی $j = 1, 2, \ldots, l$ باشد یعنی $Range(\Gamma) = \{A_1, A_2, \ldots, A_l\}$، آن‌گاه نظیر هر A_j مجموعه $\Gamma^{-1}(A_j) \subset X$ به‌صورت زیر تعریف می‌شود

$$\Gamma x = A_j \subset \Omega \longrightarrow \Gamma^{-1}(A_j) = \{x | x \in X, \Gamma x = A_j\} \qquad (۱٦)$$

پس نگاشت Γ فضای X را به‌صورت زیر افراز می‌کند

$$\{\Gamma^{-1}(A_1), \Gamma^{-1}(A_2), \ldots, \Gamma^{-1}(A_l), \Gamma^{-1}(\emptyset)\} \qquad (۱۷)$$

که $\Gamma^{-1}(\emptyset) = \{x | x \in X, \Gamma x = \emptyset\}$ مجموعه اعضایی از X است که در دامنه Γ قرار ندارند. حال به هر مجموعه در افراز (۱۷) مقدار احتمال آن نظیر می‌شود

$$\{\mu(\Gamma^{-1}(A_1)), \mu(\Gamma^{-1}(A_2)), \ldots, \mu(\Gamma^{-1}(A_l)), \mu(\Gamma^{-1}(\emptyset))\} \qquad (۱۸)$$

[۱۷] Set of compatible probability distributions
[۱۸] Focal elements

با توجه به تعریف افراز واضح است که مجموع احتمال‌ها در (۱۸) برابر با یک است. A_j ها را اعضای کانونی [18] می‌نامیم و جرم [19] نظیر هر عضو کانونی A_j به صورت زیر تعریف می‌شود

$$m_j = m(A_j) = \frac{\mu(\Gamma^{-1}(A_j))}{1-\mu(\Gamma^{-1}(\emptyset))}, \ \forall A_j \in Range(\Gamma) \tag{19}$$

واضح است که $\sum m_j = 1$، و به‌سادگی می‌توان دید که

$$\begin{cases} T^* = \bigcup_{j:A_j \cap T \neq \emptyset} \Gamma^{-1}(A_j) \\ T_* = \bigcup_{j:A_j \subset T} \Gamma^{-1}(A_j) \end{cases}$$

$$\Longrightarrow \begin{cases} \mu(T^*) = \sum_{j:A_j \cap T \neq \emptyset} \mu(\Gamma^{-1}(A_j)) \\ \mu(T_*) = \sum_{j:A_j \subset T} \mu(\Gamma^{-1}(A_j)) \end{cases} \tag{20}$$

و با توجه به این که $\mu(\Omega^*) = \mu(\Omega_*) = 1 - \mu(\Gamma^{-1}(\emptyset))$؛ از روابط (۱۳) و (۱۹) و (۲۰) نتیجه می‌شود

$$\begin{cases} \mu_h(T) = \frac{\mu(T^*)}{\mu(\Omega^*)} = \sum_{j:A_j \cap T \neq \emptyset} m_j, \forall T \subset \Omega \\ \mu_l(T) = \frac{\mu(T_*)}{\mu(\Omega^*)} = \sum_{j:A_j \subset T} m_j, \forall T \subset \Omega \end{cases} \tag{21}$$

یعنی حدود پایین و بالای احتمال برای یک پیشامد T روی Ω را می‌توان تنها با داشتن اعضای کانونی و جرم نظیر آنها از رابطه‌های ۱۲ و ۱۳ محاسبه کرد. در این حالت دیگر مانند ۱۴ نیاز مستقیم به Γ، μ و فضای X نداریم. علاوه بر خواص پیش گفته حدود پایین و بالای احتمال، مذکور در (۳) تا (۸) و (۱۴)، به سادگی می‌توان دید که

$$\mu_l(T) + \mu_h(T^c) = 1, \ \forall T \subset \Omega$$

مثال ۲ مثال ۱ را در نظر بگیرید، بنا به جدول ۱ نگاشت T را به‌صورت زیر داریم

$$\Gamma x_1 = \{w_1\}, \Gamma x_2 = \{w_1, w_2\}, \Gamma x_3 = \{w_3\},$$

$$\Gamma x_4 = \{w_2, w_3\}, \Gamma x_5 = \{w_1, w_2\}.$$

فرض کنید بخواهیم حدود پایین و بالای احتمال پیشامد این که آیا شرکت در وضعیت ورشکستگی (w_1) یا سود دهی پایدار(w_3) قرار دارد را بیابیم، یعنی پیشامد $T = \{w_1, w_3\}$. بنا به ۹ و ۱۰ داریم

$$T_* = \{x_1, x_3\}, \quad T^* = \{x_1, x_2, x_3, x_4, x_5\}$$

با توجه به جدول ۲

$$\mu(T_*) = 0/8, \ \mu(T^*) = 1$$

و چون $\mu(\Gamma^*) = 1$، آن‌گاه از ۱۲ داریم

$$\mu_1(T) = 0/8, \ \mu_h(T) = 1$$

و لذا

$$P(T) \in [0/8, 1]$$

در جدول ۳ مقادیر حدود پایین و بالای احتمال برای پیشامدهای مختلف درج شده است

جدول ۳. مقادیر حدود پایین و بالای احتمال مثال ۱

پیشامد	حد پایین احتمال	حد بالای احتمال
$\{w_1\}$	۰/۷	۰/۹
$\{w_2\}$	۰	۰/۲
$\{w_3\}$	۰/۱	۰/۱
$\{w_1, w_2\}$	۰/۹	۰/۹
$\{w_1, w_3\}$	۰/۸	۱
$\{w_2, w_3\}$	۰	۰/۳

[19] Mass

۳ نظریه شواهد شفر

نظریه شواهد شفر [۴] یک توسعه طبیعی نظریه احتمال بیزی است. خواهیم دید این نظریه به نتایجی هم‌ارز با نظریه حدود پایین و بالا احتمال دمپستر، تنها با تعبیری متفاوت، منجر می‌شود.

تعریف ۱ تابع $m : P_\Omega \longrightarrow [\circ, \text{۱}]$ را یک تابع جرم[۲۰] گوییم هر گاه $m(\emptyset) = \circ$ و $\sum_{A\subset\Omega} m(A) = \text{۱}$.

هر $A \subset \Omega$ که برای آن $m(A) > \circ$، عضو کانونی تابع جرم m نامیده می‌شود. اجتماع تمامی عضوهای کانونی m را هسته m می‌نامیم. به مجموعه اعضای کانونی همراه با جرم آنها یک ساختار اعتقادی گفته می‌شود که دانستن آن معادل با دانستن تابع جرم m است.

تعریف ۲ تابع $Bel : P_\Omega \longrightarrow [\circ, \text{۱}]$ را یک تابع باور[۲۱] (اعتقاد) گوییم هر گاه $Bel(\emptyset) = \circ$، $Bel(\Omega) = \text{۱}$ و

$$\begin{aligned} &Bel(B_\text{۱} \cup B_\text{۲} \cup \ldots \cup B_k) \\ &\geq \sum_{\emptyset \neq I = \{\text{۱},\text{۲},\ldots,k\}} (-\text{۱})^{|I|+\text{۱}} Bel(\cap_{i\in I} B_i), B_i \subset \Omega \end{aligned} \qquad (\text{۲۲})$$

قضیه ۱ اگر m یک تابع جرم باشد، آن‌گاه تابع Bel با تعریف

$$Bel(B) = \sum_{j:A_j\subset B} m_j \;,\; \forall B \subset \Omega \qquad (\text{۲۳})$$

یک تابع باور است، و بر عکس اگر Bel یک تابع باور باشد، آن‌گاه تابع m با تعریف

$$m(B) = \sum_{A\subset B} (-\text{۱})^{|B-A|} Bel(A) \;,\; \forall B \subset \Omega \qquad (\text{۲۴})$$

یک تابع جرم است.

برای اثبات می‌توان به [۴] یا [۵] مراجعه کرد. اثبات قسمت دوم قضیه با استفاده از مفهوم وارون موبیوس[۲۲] انجام می‌شود. با توجه به قضیه ۱ برای بنا کردن نظریه شواهد می‌توان اول تابع جرم را تعریف کرد و سپس تابع باور را نتیجه گرفت و یا بر عکس ابتدا تابع باور را تعریف کرد و سپس تابع جرم را نتیجه گرفت. به لحاظ تاریخی، شفر ابتدا تابع باور را معرفی کرد، ولی به نظر می‌آید روش اول به لحاظ محاسباتی برتری‌هایی داشته باشد [٦]. به این ترتیب که به جای آن که با تعداد زیادی نامساوی سر و کار داشته باشیم، با معدودی عضو کانونی و جرم آنها سر و کار داریم که محاسبات و نتیجه‌گیری‌ها را ساده‌تر می‌کند. دیده می‌شود که تابع Bel در واقع همان حد پایین احتمال، μ_l، است. این نکته در ۲۱ به روشنی قابل درک است. اما توجه به این نکته حائز اهمیت است که بیان اخیرالذکر از نظریه شواهد، به لحاظ پایه‌های نظری، کاملا مستقل از نظریه دمپستر است. در واقع، هیچ ارجاعی به حدی برای احتمال بودن تابع باور نداده‌ایم، بلکه آن را در یک ساختار اصل موضوعی بنا کرده‌ایم. تعبیر Bel به عنوان حد پایین احتمال چیزی نیست که از ساختار نظریه شواهد شفر نتیجه شود بلکه همسویی نتایج با نظریه حدود پایین و بالا احتمال دمپستر ما را به این نتیجه می‌رساند. در واقع دو خاستگاه مختلف به نتایج یکسانی رسیده‌اند. پس برای این نتایج می‌توان هر دو معنا را قائل بود، یعنی Bel هم تابع باور است و هم حد پایین احتمال و m هم تابع جرم است و هم جرم احتمال پیشامدهای حاصل از افراز X به سبب نگاشت چند مقداری Γ.

در ادامه تابع موجه‌نمایی[۲۳] را تعریف می‌کنیم که تعریفی هم‌ارز با حد بالای احتمال، μ_h، دارد.

[۲۰] Mass function
[۲۱] Belief function
[۲۲] Mobious Inversion
[۲۳] Plausibility function

تعریف ۳ تابع $Pl : P_\Omega \longrightarrow [0,1]$ را یک تابع موجه‌نمایی گوییم هر گاه $Pl(\emptyset) = 0$، $Pl(\Omega) = 1$ و

$$Pl(B_1 \cap B_2 \cap \ldots \cap B_k) \le \sum_{\emptyset \neq I = \{1,2,\ldots,k\}} (-1)^{|I|+1} Pl(\cup_{i\in I} B_i)\ ,\ B_i \subseteq \Omega \qquad (۲۵)$$

قضیه ۲ اگر Bel یک تابع باور باشد آن‌گاه تابع Pl با تعریف

$$Pl(B) = 1 - Bel(B^c), \quad \forall B \subset \Omega \qquad (۲۶)$$

یک تابع موجه‌نمایی است. هم چنین اگر m تابع جرم نظیر Bel باشد، آن‌گاه

$$Pl(B) = \sum_{j: A_j \cap B \neq \emptyset} m_j\ ,\ \forall B \subset \Omega \qquad (۲۷)$$

باتوجه به قضیه‌های ۱ و ۲ مشاهده می‌شود که تابع موجه‌نمایی همان میزان آگاهی را در بردارد که تابع باور و یا تابع جرم در بر دارند، بدین ترتیب که با در دست داشتن هر یک از آنها می‌توان دو تابع دیگر را ساخت. خواص Bel و Pl هم‌ارز خواص μ_l و μ_h است که در بخش ۲ بررسی شد. شکل ۲ نحوه محاسبه Bel و Pl یک پیشامد را تشریح می‌کند. به این ترتیب است که نظریه دمپستر—شفر شکل می‌گیرد. وقتی از نظریه دمپستر—شفر صحبت می‌شود یعنی مجموعه نظریه حدود پایین و بالا احتمال دمپستر و نظریه شواهد شفر، همراه با بحث معادل بودن آنها و تمامی نتایج و تعابیری که از دو سو به‌دست می آید.

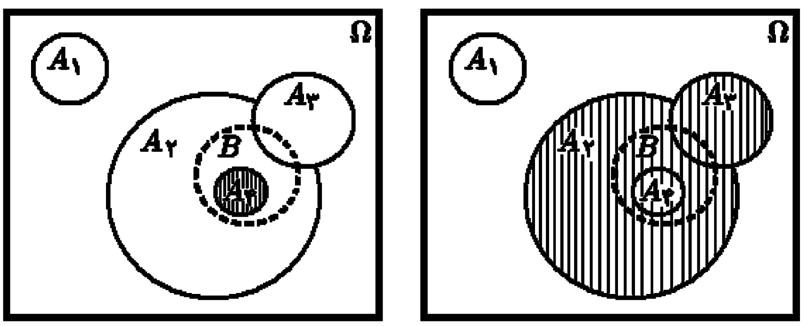
در ادامه به معرفی دو ساختار اعتقادی ویژه می‌پردازیم.

قضیه ۳ فرض کنید Bel یک تابع باور باشد. شرط لازم و کافی برای آن که Bel یک تابع احتمال باشد، آن است که

$$Bel(B) = Pl(B) \quad \forall B \subset \Omega \qquad (۲۸)$$

یعنی در این حالت خاص هیچ نوع نایقینی در بیان احتمال پیشامد نداریم و طول بازه عدم آگاهی، $Pl(B) - Bel(B)$، برابر با صفر است. می‌توان دید که ۲۸ تنها وقتی برآورده می‌شود که اعضای کانونی به صورت زیرمجموعه‌های تک عضوی Ω درآیند. در این حالت تابع جرم همان تعبیر تابع توزیع احتمال را می‌یابد. پس نظریه احتمال حالت خاص نظریه دمپستر—شفر است.

در این مقاله Ω را مجموعه‌ای با تعداد اعضای متناهی، n عضو، در نظر گرفته‌ایم. پس تعداد اعضای کانونی، که معادل با تعداد اعضای برد Γ یا تعداد اعضای دامنه m است، حداکثر $2^n - 1$ است که شامل تمامی زیرمجموعه‌های Ω به جز $\emptyset$ می‌شود. حال به حالت خاص دوم می‌پردازیم.



$Bel(B) = m(A_4)$ $\qquad$ $Pl(B) = m(A_2) + m(A_3) + m(A_4)$

شکل ۲. شیوه محاسبه توابع باور و موجه‌نمایی با داشتن اعضای کانونی و جرم نظیر آنها

قضیه ۴ ساختاری اعتقادی را در نظر بگیرید که تنها شامل یک عضو کانونی و آن هم Ω باشد، یعنی

$$m(\Omega) = 1 \quad , \quad m(B) = 0 \quad \forall B \subset \Omega \; (B \neq \Omega) \qquad (29)$$

آنگاه داریم

$$Bel(B) = 0 \quad Pl(B) = 1 \quad \forall B \subset \Omega \; (B \neq \emptyset, \Omega) \qquad (30)$$

یعنی در این حالت خاص بیشترین میزان نایقینی ممکن در بیان احتمال پیشامد را داریم و طول بازه عدم آگاهی برای تمامی پیشامدها، جز Ω و $\emptyset$، برابر با حداکثر ممکن یعنی یک است. این حالت را عدم آگاهی کامل[24] می‌نامیم. تعبیر عدم آگاهی کامل حلال یکی از مشکلات نظریه بیز است: دیگر لازم نیست عدم اطلاعات را هم‌ارز با توزیع یکنواخت به عنوان مدل احتمال پیشین بدانیم، بلکه می‌توانیم به جای آن تعبیر عدم آگاهی کامل مطابق با قضیه ۴ را جایگزین کنیم.

مثال ۳ دوباره مثال ۱ را در نظر بگیرید و فرض کنید بخواهیم مقادیر باور و موجه‌نمایی پیشامد T را به‌دست آوریم. بنا به مثال ۲، تابع جرم به راحتی محاسبه می‌شود. می‌توان دید که برد نگاشت $\{A_1 = \{w_1\}, A_2 = \{w_1, w_2\}, A_3 = \{w_3\}, A_4 = \{w_2, w_3\}\}$ است، بنا به ۱۶ نگاشت Γ افراز زیر از فضای X را ایجاد می کند

$$\begin{aligned} \Gamma^{-1}(A_1) &= \{x_1\} \\ \Gamma^{-1}(A_1) &= \{x_2, x_5\} \\ \Gamma^{-1}(A_1) &= \{x_3\} \\ \Gamma^{-1}(A_1) &= \{x_4\} \end{aligned}$$

لذا، طبق ۱۹، تابع جرم به صورت زیر محاسبه می‌شود

$$m(A_1) = 0/7, \; m(A_2) = 0/2,$$
$$m(A_3) = 0/1, \; m(A_4) = 0$$

لذا اعضای کانونی ساختار اعتقادی، $\{A_1, A_2, A_3\}$ است، و لذا برای پیشامد $T = \{w_1, w_2\}$، از طریق روابط ۲۳ و ۲۷، مقادیر باور و موجه‌نمایی عبارت‌اند از

$$Bel(T) = m(A_1) + m(A_3) = 0/8$$
$$Pl(T) = m(A_1) + m(A_2) + m(A_3) = 1$$

برای سایر پیشامدها، مقادیر جدول ۳ برقرار است.

۴ قاعده اعمال شرط دمپستر و قاعده ترکیب دمپستر

یک ساختار اعتقادی در نظریه دمپستر—شفر را به‌صورت زیر در نظر می‌گیریم

$$\{A_1, A_2, \ldots, A_l\} \quad A_j \subset \Omega \; (A_j \neq \emptyset) \qquad (31)$$

$$\{m(A_1), m(A_2), \ldots, m(A_l)\} \quad m(A_j) > 0,$$
$$\sum m(A_j) = 1 \qquad (32)$$

که در آن A_jها اعضای کانونی و $m(A_j)$ها مقادیر جرم نظیر هستند. حال فرض کنید شرط R روی فضای Ω داده شده است، می‌خواهیم احتمال پیشامد T را به شرط R بیابیم. تعبیر دمپستر از مسئله چنین است که اندازه احتمال μ روی فضای X تغییر نکرده و تنها مجموعه $\Omega - R$ ناممکن و در نتیجه نامحتمل شده است. پس کافی است که برد Ω را محدود به R کنیم. در این حالت $\Gamma'(x) = \Gamma(x) \cap R$ نگاشت جدید بوده و برد آن به صورت زیر تحدید می‌شود

$$\{A_1 \cap R, A_2 \cap R, \ldots, A_l \cap R\} \qquad (33)$$

[24] Total ignorance

اما تنها اگر $A_j \cap R \neq \emptyset$ می‌توان جرم $m(A_j)$ را به آن تخصیص داد. یعنی آن دسته اعضای کانونی A_j که به تمامی خارج از شرط R قرار می‌گیرند پس از تحدید به R دیگر نمی‌توانند عضو کانونی باشند. برای رفع مشکل، جرم واقع شده بیرون R به صورت یکنواخت بین اعضا کانونی جدید توزیع می‌شود، یعنی

$$m'(A_j \cap R) = \begin{cases} \frac{1}{k} m(A_j) & A_j \cap R \neq \emptyset \\ ۰ & A_j \cap R = \emptyset \end{cases}$$
$$m(A_j) > ۰, \sum m(A_j) = ۱ \qquad (۳۴)$$

به این ترتیب پس از اعمال شرط، نگاشت جدید Γ' و تابع جرم جدید m' برای اعضای کانونی به فرم $A_j \cap R(\neq \emptyset)$ به‌دست می‌آید. شکل ۳ تعبیر اعمال شرط را به خوبی نشان می‌دهد.

مثال ۴ ساختار اعتقادی مثال ۱ را در نظر بگیرید. فرض کنید اکنون اطلاعات جدیدی به‌دست آمده است که وضعیت سوددهی شرکت در یکی از دو حالت ورشکستگی ($w_۱$) و یا سوددهی ناپایدار ($w_۲$) قرار دارد، در نتیجه می‌خواهیم تابع جرم جدید را مشروط به $R = \{w_۱, w_۲\}$ به‌دست آوریم. بنا به ۳۴ داریم

$$m'(A_۱ \cap R) = m'(\{w_۱\}) = (۱/k).m(A_۱) = (۱/k).(۰/۷)$$
$$m'(A_۲ \cap R) = m'(\{w_۱, w_۲\}) = (۱/k).m(A_۲) = (۱/k)(۰/۲)$$

که در آن ضریب بهنجارسازی برابر است با

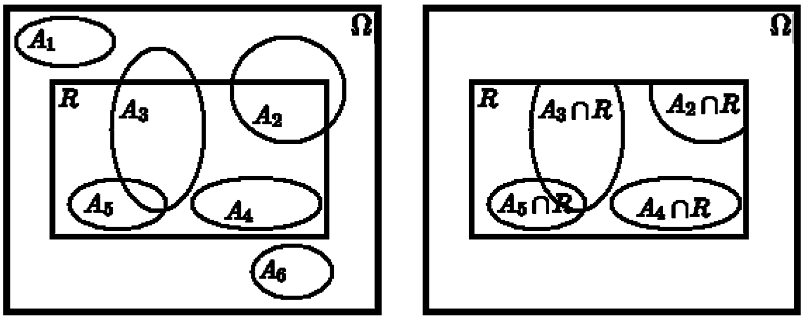
$$k = ۰/۲ + ۰/۷ = ۰/۹$$

لذا ساختار اعتقادی جدید به صورت زیر می‌باشد

$$\{B_۱ = \{w_۱\}, B_۲ = \{w_۱, w_۲\}\}$$
$$\{m'(B_۱) = ۰/۷۸, m'(B_۲) = ۰/۲۲\}$$

توسعه دیدگاه فوق به قاعده ترکیب دمپستر منجر می‌شود. فرض کنید روی فضای X_i اندازه احتمال $\mu_i,\ i = ۱,۲$ را داشته باشیم. نیز فرض کنید فضاهای X_i مجزا و اندازه‌های احتمال μ_i مستقل از یکدیگر باشند. نظیر هر فضای X_i نگاشت $\Gamma_i : X_i \longrightarrow P_\Omega$ را در نظر می‌گیریم. دیدیم که هر Γ_i نظیر μ_i یک ساختار حدود پایین و بالای احتمال روی Ω ایجاد می‌کند. حال سوال این است که اثر مجموع اطلاعات ما که حاصل از منابع مجزای X_i و توابع احتمال مستقل μ_i است چه می‌تواند باشد؟ با توجه به فرض منابع مستقل، ترکیب منابع ساده است

$$X = X_۱ \times X_۲, \quad \mu = \mu_۱ \mu_۲ \qquad (۳۵)$$
$$\Gamma(x) = \Gamma_۱(x_۱) \cap \Gamma_۲(x_۲)$$
$$\Gamma : X \to \Omega \to x = (x_۱, x_۲) \in X \qquad (۳۶)$$



شکل ۳. در شکل سمت چپ اعضای کانونی A_j و شرط R نشان داده شده است. شکل سمت راست اعضای کانونی به فرم $A_h \cap R(\neq \varnothing)$ بعد از اعمال شرط را نشان می‌دهد.

نگاشت $\Gamma : X_{۱} \times X_{۲} \longrightarrow \Omega$ همراه با فضای احتمال $\mu = \mu_{۱}\mu_{۲}$ حدود بالا و پایین احتمال را روی Ω ایجاد می‌کند. تعریف Γ به صورت رابطه ۳٦ یعنی x تنها در صورتی با w سازگار است که هر دوی $x_{۱}$ و $x_{۲}$ با w سازگار باشند. تعریف ۳٦ ربطی به فرض استقلال منابع ندارد و تنها یک نوع ترکیب عطفی روابط سازگاری است. این روش برای ترکیب آگاهی‌های به دست آمده از دو منبع مستقل به سادگی قابل توسعه به بیش از دو منبع است. این شیوه‌ی ترکیب را می‌توان برحسب توابع جرم نیز بیان کرد. توابع جرم نظیر $\Gamma_{۱}$ و $\Gamma_{۲}$ را در نظر می‌گیریم

$$Range(\Gamma_{۱}) = \{A_{۱}, A_{۲}, \dots, A_l\} \quad m_{۱}(A_i), i = ۱, ۲, \dots, l \qquad (۳۷)$$
$$Range(\Gamma_{۲}) = \{B_{۱}, B_{۲}, \dots, B_k\} \quad m_{۲}(B_j), j = ۱, ۲, \dots, k \qquad (۳۸)$$

با توجه به۳۵، ۳٦، ۳۷ و ۳۸ خواهیم داشت

$$Range(\Gamma) = \{A_i \cap B_j \mid A_i \cap B_j \neq \emptyset, i = ۱, ۲, \dots, l, j = ۱, ۲, \dots, k\} \quad (۳۹)$$

حال اگر $C \in Range(\Gamma)$، با توجه به ۳۵ و ۳۹، تابع جرم نظیر Γ عبارت خواهد بود از

$$m(C) = \frac{۱}{N} \sum_{A_i \cap B_j = C} m_{۱}(A_i) m_{۲}(B_j) \qquad (۴۰)$$

که در آن N ضریب بهنجار کردن تابع جرم m است، و عبارت است از

$$N = \sum_{A_i \cap B_j \neq \emptyset} m_{۱}(A_i) m_{۲}(B_j) = ۱ - \sum_{A_i \cap B_j = \emptyset} m_{۱}(A_{۱}) m_{۲}(B_j) \quad (۴۱)$$

رابطه ۴۰ موسوم به قاعده ترکیب دمپستر است که کاربرد گسترده‌ای در مسائل ترکیب اطلاعات یافته است. این شیوه‌ی ترکیب در عین سادگی و زیبایی می‌تواند بحث انگیز باشد، چرا که فرض منابع اطلاعاتی مستقل اگر چه تعریف ریاضی واضحی دارد، ولی تحقق آن در عمل غالبا بعید است. در این زمینه، بررسی دو حالت خاص از ترکیب اطلاعات حاصل از منابع مستقل جالب توجه است. ساختاری اعتقادی را در نظر بگیرید که تنها شامل یک عضو کانونی R باشد، یعنی

$$m(R) = ۱, m(B) = ۰ \quad \forall B \subset \Omega\ (B \neq R) \qquad (۴۲)$$

به سادگی می‌توان دید که حاصل ترکیب ۴۲ با هر ساختار دیگر به منزله اعمال شرط R به آن ساختار اعتقادی است. پس قاعده اعمال شرط دمپستر حالت خاص قاعده ترکیب دمپستر است. حال در حالت خاص $R = \Omega$ ترکیب ۴۲ با هر ساختار اعتقادی دیگر به منزله اعمال شرط Ω است که هیچ تغییری در ساختار اعتقادی ایجاد نخواهد کرد. پس ساختار اعتقادی از نوع عدم آگاهی کامل، ۲۹، عضو خنثی عمل ترکیب ساختارهای اعتقادی به کمک قاعده ترکیب دمپستر است. یعنی چنین منبعی عملا حاوی هیچ اطلاعاتی نیست. این تعبیر دیگری برای عدم آگاهی کامل است. در ترکیب ساختارهای اعتقادی به کمک قاعده ترکیب دمپستر، اضافه شدن هر آگاهی جدید باعث کاهش تعداد اعضای موجود در عضوهای کانونی می شود و در واقع به سمت تابع احتمال حرکت می‌کنیم و این موافق با دیدگاه کلی ما می‌باشد که کم شدن نایقینی همگام با افزایش اطلاعات است.

مثال ۵ فرض کنید نگاشت‌های $\Gamma_{۱}$ و $\Gamma_{۲}$ از $X_{۱}$ و $X_{۲}$ با اندازه‌های احتمال $\mu_{۱}$ و $\mu_{۲}$ از دو منبع اطلاعاتی در دسترس است. منبع اول را همانند جداول ۱ و ۲ در مثال ۱ در نظر بگیرید و برای منبع دوم فرض کنید مقادیر مندرج در جداول ۴ و ۵ داده شده است

حال همانند مثال ۲، نگاشت دوم را به‌صورت زیر می‌نویسیم

$$\Gamma_{۲} x_{۱} = \{w_{۲}\}, \Gamma_{۲} x_{۲} = \{w_{۱}, w_{۲}\},$$
$$\Gamma_{۲} x_{۳} = \{w_{۱}, w_{۳}\}, \Gamma_{۲} x_{۴} = \{w_{۲}\}, \Gamma_{۲} x_{۵} = \{w_{۳}\}$$

جدول ۶ برد نگاشت $\Gamma(x) = \Gamma_1(x_1) \cap \Gamma_2(x_2)$ را نشان می‌دهد

جدول ۴. رابطه (γ) وضعیت سوددهی و میزان تولید منبع دوم (۱ به معنی وجود رابطه و ۰ به معنی عدم وجود آن است)

رابطه γ	x_1	x_2	x_3	x_4	x_5
w_1	۰	۱	۱	۰	۰
w_2	۱	۱	۰	۱	۰
w_3	۰	۰	۱	۰	۱

جدول ۵. اندازه احتمال روی فضای X_2

توزیع μ	x_1	x_2	x_3	x_4	x_5
$\mu(x_i)$	۰/۱۵	۰	۰/۷	۰/۱۵	۰

جدول ۶. نگاشت $\Gamma(x) = \Gamma_1(x_1) \cap \Gamma_2(x_2)$

$X = X_1 \times X_2$	x_1	x_2	x_3	x_4	x_5
x_1	$\varnothing$	$\{w_1\}$	$\{w_1\}$	$\varnothing$	$\varnothing$
x_2	$\{w_2\}$	$\{w_1, w_2\}$	$\{w_1\}$	$\{w_2\}$	$\varnothing$
x_3	$\varnothing$	$\varnothing$	$\{w_3\}$	$\varnothing$	$\{w_3\}$
x_4	$\{w_2\}$	$\{w_2\}$	$\{w_3\}$	$\{w_2\}$	$\{w_3\}$
x_5	$\{w_2\}$	$\{w_1, w_2\}$	$\{w_1\}$	$\{w_2\}$	$\varnothing$

جدول ۷. اندازه $\mu = \mu_1\mu_2$ روی فضای $X = X_1 \times X_2$

$X = X_1 \times X_2$	x_1	x_2	x_3	x_4	x_5
x_1	۰/۱۰۵	۰	۰/۴۹	۰/۱۰۵	۰
x_2	۰/۰۳	۰	۰/۱۴	۰/۰۳	۰
x_3	۰/۱۰۵	۰	۰/۰۷	۰/۱۰۵	۰
x_4	۰	۰	۰	۰	۰
x_5	۰	۰	۰	۰	۰

در جدول ۷ اندازه $\mu = \mu_1\mu_2$ روی فضای $X = X_1 \times X_2$ نشان داده شده است. با محاسبات لازم، تابع جرم متناظر با Γ و μ، از جداول ۶ و ۷، به‌ترتیب زیر به‌دست می آید

$$m(\{w_1\}) = ۰/۶۳/(۱ - ۰/۲۴),$$

$$m(\{w_2\}) = ۰/۰۶/(۱ - ۰/۲۴),$$

$$m(\{w_3\}) = ۰/۰۷/(۱ - ۰/۲۴)$$

حال همین نتیجه را از ترکیب مستقیم دو ساختار اعتقادی به کمک قاعده ترکیب دمپستر به‌دست می آوریم. تابع جرم به‌دست آمده در مثال ۳ را در نظر بگیرید

$$m_1(A_1) = ۰/۷,\ m_1(A_2) = ۰/۲,\ m_1(A_3) = ۰/۱$$

هم چنین تابع جرم زیر از اطلاعات منبع دوم استخراج می‌شود

$$\{B_1 = \{w_2\}, B_2 = \{w_1, w_3\}\}$$

$$\{m_2(B_1) = ۰/۳, m_2(B_2) = ۰/۷\}$$

اگر این فرض برقرار باشد که دو ساختار اعتقادی مذکور از منابعی مستقل حاصل شده‌اند، قاعده ترکیب دمپستر را می‌توان برای تعیین ساختار اعتقادی حاصل از ترکیب اطلاعات دو منبع به‌کار برد. بدین ترتیب روابط ۴۰ و ۴۱ به‌دست می‌دهد

$$m(\{w_1\}) = \frac{1}{N}۰/۶۳, m(\{w_2\}) = \frac{1}{N}۰/۰۶$$

$$m = (\{w_3\}) = \frac{1}{N}۰/۰۷$$

$$N = ۱ - m(\phi) = ۱ - ۰/۲۴ = ۰/۷۶$$

$$m(\{w_1\}) = ۰/۸۳, m(\{w_2\}) = ۰/۰۸, m(\{w_3\}) = ۰/۰۹$$

یعنی ترکیب اطلاعات دو منبع بیشترین باور را به این که شرکت در وضعیت ورشکستگی قرار دارد، القا می‌کند. این نتیجه دقیقاً مطابق تابع جرمی است که در ابتدای مثال حاصل گردید، که نتیجه‌ای قابل انتظار بود.

۵ مجموعه توزیع‌های احتمال سازگار با یک ساختار اعتقادی

در بخش‌های ۲ و ۳ با اصول نظریه دمپستر–شفر، آن چنان که توسط پایه گذارانش ارائه شده بود، آشنا شدیم. درمجموع نظریه دمپستر–شفر را می‌توان توسعه نظریه احتمال دانست، به قسمی که تعبیر نایقینی و عدم آگاهی در بیان خود احتمال را نیز در بر گیرد. در این بخش اما، به نظریه شواهد

از زاویه‌ای دیگر نگاه خواهیم کرد. به این ترتیب که نظیر هر ساختار اعتقادی، مجموعه توزیع‌های احتمال سازگار با آن را تعریف می‌کنیم، و خواهیم دید که مسئله یافتن تابع باور و تابع موجه‌نمایی تبدیل به مسئله یافتن کمینه و بیشینه احتمال یک پیشامد روی مجموعه توزیع‌های احتمال سازگار با ساختار اعتقادی می‌شود. این دیدگاه نخستین بار توسط دمپستر در سال ۱۹۶۷ [۱] ارائه شد. البته این دیدگاه بعد از معرفی آن، بسیار کم مورد توجه قرار گرفته است. اکثرا ترجیح می‌دهند از روش محاسباتی و تعابیر مبتنی بر نظریه شواهد شفر استفاده کنند و اگر نیاز به تفسیر احتمالاتی باشد از تعبیر حدود پایین و بالای احتمال دمپستر استفاده کنند. در این بخش مجموعه توزیع‌های احتمال سازگار با یک ساختار اعتقادی به عنوان نمایشی معادل با ساختار اعتقادی معرفی می‌شود. این دیدگاه علاوه بر آن که پایه‌ی نظری لازم برای توسعه نظریه دمپستر—شفر به محیط‌های فازی را ایجاد می‌کند[۷]، امکان استفاده سر راست‌تر از مفاهیم موجود در نظریه غنی احتمالات را فراهم می‌سازد. نیز به وضوح نشان می‌دهد که روابط و تعاریف موجود در نظریه دمپستر—شفر، به ویژه قاعده اعمال شرط و قاعده ترکیب، تنها روش و یا لزوما طبیعی‌ترین روش تعریف چنین قواعدی نیستند. هم‌چنین به کمک این دیدگاه می‌توان تعبیر مناسبی برای مشاهده نادقیق به دست آورد. بدین ترتیب که به جای آن که اعضای کانونی را ثمره یک نگاشت چند مقداری بدانیم، آنها را دقیقا خود مشاهدات بدانیم، یعنی خروجی‌های یک تجربه تصادفی، با این تفاوت که در این تجربه تصادفی، مشاهدات به جای آن که اعضای Ω باشند زیرمجموعه‌های Ω هستند.

توزیع $P=\{p_1,p_2,\dots,p_n\}$ را روی Ω در نظر بگیرید. تابع احتمال نظیر توزیع را نیز با P نشان می‌دهیم. قصد محاسبه احتمال پیشامد T روی Ω را داریم

$$P(T)=\sum_{i:w_i\in T}p_i \qquad (43)$$

اگر توزیع احتمال P در دست بود به کمک رابطه ۴۳ محاسبه $P(T)$ سر راست بود. اما فرض ما آن است که آنچه راجع به توزیع احتمال روی Ω می‌دانیم در قالب نظریه دمپستر—شفر و به فرم یک ساختار اعتقادی، مانند ۳۱ و ۳۲، است. تابع جرم، جرم احتمال $m_j=m(A_j)$ را به عضو کانونی A_j به عنوان یک توده اطلاعاتی [۲۵] نسبت می‌دهد. می‌خواهیم جرم احتمال m_j را بین اعضای Ω تقسیم کنیم. به این منظور سهم احتمالی که از m_j به w_i می‌رسد را با m_{ij} نشان می‌دهیم. سه دسته محدودیت زیر روی m_{ij} ها بدیهی می‌نماید

$$m_{ij}\ \geq\ 0 \qquad i=1,2,\dots,n,\ j=1,2,\dots,l \qquad (44)$$

$$m_{ij}\ =\ 0 \quad if\ w_i\notin A_j \qquad \begin{array}{l} i=1,2,\dots,n\\ j=1,2,\dots,l\end{array} \qquad (45)$$

$$m_j\ =\ \sum_{i=1}^{n}m_{ij} \qquad j=1,2,\dots,l \qquad (46)$$

در این چارچوب p_iها به صورت زیر تعریف می‌شوند

$$p_i=\sum_{j=1}^{l}m_{ij} \qquad i=1,2,\dots,n \qquad (47)$$

و از مقایسه ۴۷ و ۴۳ نتیجه می‌شود

$$P(T)=\sum_{i:W_i\in T}\sum_{j=1}m_{ij}^{l} \qquad (48)$$

تعریف ۴ ساختار اعتقادی ۳۱ و ۳۲ در چارچوب نظریه دمپستر—شفر را در نظر می‌گیریم. مجموعه تمامی توزیع‌های احتمال به فرم $P=\{p_1,p_2,\dots,p_n\}$ روی Ω که در تمامی شرایط ۴۴، ۴۵، ۴۶ و ۴۷ صدق کنند را مجموعه توزیع‌های احتمال سازگار با ساختار اعتقادی می‌نامیم و با S نشان می‌دهیم.

[۲۵] Granule of information

قضیه ۵ ساختار اعتقادی ۳۱ و ۳۲ مفروض است. برای هر $T \in \Omega$، پاسخ مسئله کمینه‌سازی $P(T)$ در ۴۳ محدود به شرایط ۴۴، ۴۵، ۴۶ و ۴۷ – یعنی محدود به توزیع‌هایی که با S مشخص شده است– برابر است با $Bel(T)$. به همین ترتیب پاسخ مسئله بیشینه‌سازی $P(T)$ محدود به مجموعه توزیع‌های احتمال سازگار با ساختار اعتقادی، S، برابر است با $Pl(T)$.

برای اثبات قضیه به [۱] مراجعه شود.

مجموعه توزیع‌های احتمال سازگار باساختار اعتقادی، با تعدادی محدودیت خطی از نوع مساوی و نامساوی معرفی شده‌اند، پس یک فضای پاسخ‌های ممکن[۲۶] محدب را برای مسئله بهینه‌سازی خطی در قضیه ۵ ایجاد می‌کنند. برای حل مسئله بهینه‌سازی خطی راه حل موثری چون روش سیمپلکس موجود است. مجموعه توزیع‌های احتمال سازگار با ساختار اعتقادی از نظر محتوای اطلاعاتی معادل با تابع جرم، تابع‌باور و تابع موجه‌نمایی است. بدین معنا که با دانستن هر کدام از آنها سه دیگر را می‌توان به‌دست آورد. قضیه ۵ بار دیگر رابطه (۵۱) را تایید می‌کند که $Bel(T) \leq P(T) \leq Pl(T), \quad \forall T \subset \Omega$. در واقع در محاسبه $P(T)$ حد پایین احتمال متناظر با حالتی است که جرم احتمال تمامی اعضای کانونی دارای اشتراک ناتهی با T که زیرمجموعه T نیستند را منتسب به اعضای خارج T کنیم. در مقابل، حد بالای احتمال متناظر با حالتی است که جرم احتمال تمامی اعضای کانونی دارای اشتراک ناتهی با T که زیرمجموعه T نیستند را منتسب به اعضای داخل T کنیم. تکلیف اعضای کانونی زیرمجموعه T و نیز اعضای کانونی جدا از T که در هر حال روشن است. بنا بر این در مجموعه توزیع‌های سازگار برای هر یک از دو مسئله کمینه یا بیشینه‌سازی توزیع‌های بسیاری می‌تواند وجود داشته باشند که برای یک T خاص به $Bel(T)$ یا $Pl(T)$ منجر شوند. این درجه آزادی با توجه به تعبیر پیش گفته از دو حالت کمینه و بیشینه واضح است. باید توجه داشت که در حالت کلی برای یک T خاص $Bel(T)$ و $Pl(T)$ غالبا به ازای توزیع‌های متمایزی در S حادث می‌شوند، و نیز برای T های مختلف غالبا توزیع‌های متمایزی در S منجر به پاسخ بهینه –در کمینه یا بیشینه‌سازی– می‌شوند.

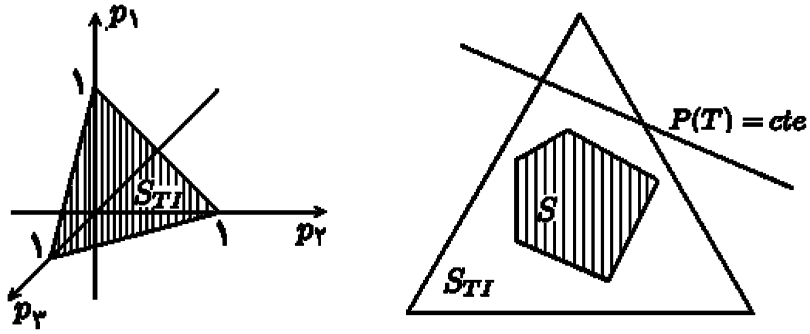
بررسی یک حالت خاص در اینجا جالب است. ساختار اعتقادی از نوع عدم آگاهی کامل، ۲۹، را در نظر بگیرید. در این حالت $l = ۱$ و $A_۱ = \Omega$. اگر محدودیت‌های ۴۴، ۴۵، ۴۶ و ۴۷ را مرور کنیم مشاهده می‌شود که رابطه ۴۷ منجر به $p_i = m_i$ می‌شود، ۴۴ منجر به $p_i \geq ۰$ می‌شود، ۴۵ در این حالت موضوعیت ندارد چرا که $A_۱ = \Omega$، و ۴۶ منجر به $\sum p_i = ۱$ می‌شود. پس مجموعه توزیع‌های سازگار با عدم آگاهی کامل عبارت است از

$$S_{TI} = \{(p_۱, p_۲, \dots, p_n) | \sum_{i=۱}^{n} p_i = ۱, p_i \geq ۰, i = ۱, ۲, \dots, n\} \quad (۴۹)$$

در واقع S_{TI} در ۴۹ شامل تمام توزیع‌های ممکن روی Ω می‌شود و هیچ محدودیت خاصی جز شرایط توزیع احتمال را شامل نمی‌شود. واضح است که مجموعه توزیع‌های احتمال سازگار با هر ساختار اعتقادی دیگر، زیرمجموعه‌ای از S_{TI} خواهدبود. در شکل ۴، S_{TI} و نیز یک S نمونه‌ای در حالت $n = ۳$ نشان داده شده است.

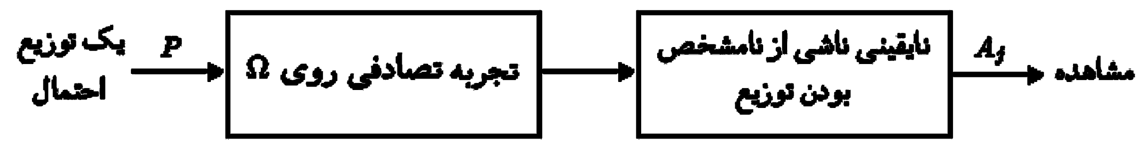
---

[۲۶] Feasible solutions space



شکل ۴. شکل سمت چپ S_{IT} را در حالت $n = 3$ روی فضای توزیع‌های احتمال (p_1, p_2, p_3) نشان می‌دهد. شکل سمت راست در صفحه S_{IT} یک S نمونه‌ای را نشان می‌دهد. خط $P(T) = cte$ در صفحه S_{IT} کمینه و بیشینه‌سازی $P(T)$ را تبیین می‌کند.

اکنون به کمک تعبیر مجموعه توزیع‌های احتمال سازگار می‌توان به یک تعبیر عینی[27] از نظریه دمپستر—شفر رسید که در آن اعضای کانونی را عینا مشاهدات یعنی خروجی‌های یک تجربه تصادفی در نظر می‌گیریم. شکل ۵ این تعبیر را روشن می‌کند.



شکل ۵. تجربه تصادفی روی Ω تحت توزیع P انجام شده اما مشاهده درگیر نوعی نایقینی از جنس نامشخص بودن به سبب نامشخص بودن توزیع P است. تنها می‌دانیم $P \in S$ و این نایقینی به مشاهده منتقل می‌شود. آنچه که مشاهده می‌شود نه عضوی از Ω بلکه زیرمجموعه‌ای از Ω است.

پس با این تعبیر اعضای کانونی، خروجی‌های یک تجربه تصادفی هستند و برای برآورد مقادیر جرم آنها می‌توان از فراوانی نسبی استفاده کرد. پس به این ترتیب به یک تعبیر مبتنی بر فراوانی[28] از احتمال در نظریه دمپستر—شفر رسیده‌ایم. یک دستاورد اساسی تعبیر مجموعه توزیع های احتمال سازگار با ساختار اعتقادی، انتقال نایقینی از نوع نامشخص بودن[29] از فضای مشاهدات Ω به فضای توزیع‌های احتمال ممکن روی Ω است. این ویژگی عملا تمامی مفاهیم، ابزارها و روش‌های توسعه یافته برای احتمال روی اعضای Ω را در دسترس نظریه دمپستر—شفر قرار می‌دهد. مثلا برای محاسبه امید ریاضی یک متغیر تصادفی W که مقادیر خود را از Ω می‌گیرد، داریم

$$E_{DST}(W) = \{E_P(W) \mid P = (p_1, p_2, \ldots, p_n) \in S\} \quad (50)$$

که در آن $E_{DST}(W)$ توسعه یافته تعبیر امید ریاضی متغیر تصادفی W در نظریه دمپستر—شفر است، و $E_P(W)$ امید ریاضی با تعریف معمول با فرض توزیع $P = (p_1, p_2, \ldots, p_n)$ است. توزیع P هر یک اعضای مجموعه توزیع‌های احتمال سازگار با ساختار اعتقادی، S، می‌تواند باشد.

مثال ۶ فرض کنید بخواهیم مجموعه توزیع‌های سازگار با ساختار اعتقادی بحث شده در مثال‌های قبلی را به‌دست

[27] Objective
[28] Frequentist view
[29] Nonspecificity

آوریم. یعنی هدف یافتن توزیع‌های $P=\{p_۱,p_۲,p_۳\}$ روی وضعیت سوددهی شرکت است که در آن p_i چگالی احتمال متناظر با w_i است. بنا به محدودیت‌های ۴۴، ۴۵ و ۴۶ داریم

$$m_{۱۱}=۰/۷,\ m_{۲۱}=۰,\ m_{۳۱}=۰$$

$$m_{۱۲}+m_{۲۲}=۰/۲,\ m_{۳۲}=۰$$

$$m_{۱۳}=۰,\ m_{۲۳}=۰,\ m_{۳۳}=۰/۱$$

و بنا به ۴۷

$$p_۱=۰/۷+m_{۱۲}$$

$$p_۲=m_{۲۲}$$

$$p_۳=۰/۱$$

لذا مجموعه توزیع‌های احتمال سازگار، با توجه محدویت‌های زیر تعیین می‌شود

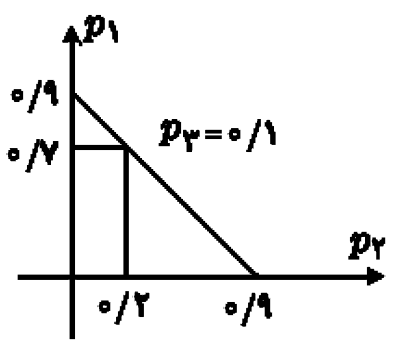
$$۰/۷\le p_۱\le ۰/۹$$

$$۰\le p_۲\le ۰/۲$$

$$p_۳=۰/۱$$

$$p_۱+p_۲=۰/۹$$

گر چه در این مثال ساده توانستیم به طور تحلیلی محدودیت‌ها را بر حسب p_iها بیان کنیم، اما در حالت کلی برای محاسبه احتمال یک پیشامد نیازی به این کار نیست و روابط قبلی می‌توانند به طور مستقیم استفاده شوند. شکل زیر مجموعه توزیع‌های ممکن را در صفحه $p_۳=۰/۱$ نشان می‌دهد



حال فرض کنید قصد داریم احتمال پیشامد $T=\{w_۱,w_۳\}$ را تعیین کنیم

$$P(T)=p_۱+p_۳=۰/۸+m_{۱۲}$$

هم چنین داریم

$$m_{۱۲}+m_{۲۲}=۰/۲,\ m_{۱۲}\ge ۰, m_{۲۲}\ge ۰$$

لذا

$$\min_P P(T)=۰/۸+۰=Bel(T)$$

$$\max_P P(T)=۰/۸+۰.۲=Pl(T)$$

۶ نکات نهایی و جمع بندی

در این مختصر به مرور مبانی نظریه دمپستر–شفر پرداختیم. سه دیدگاه متفاوت که نهایتا به نتایج یکسانی می‌رسیدند مورد بحث قرار گرفت. نظریه حدود پایین و بالای احتمال دمپستر، نظریه شواهد شفر و تعبیر مجموعه توزیع‌های احتمال سازگار با ساختار اعتقادی، که هر سه راه کارهایی معادل برای بیان و مدیریت نایقینی در احتمال را به‌دست می‌دهند، به اجمال مرور شد. در این بین روش‌های محاسباتی برآمده از نظریه شواهد شفر محبوب‌ترند، اما نگارندگان بر آن‌اند که نظریه حدود پایین و بالای احتمال دمپستر و خصوصا تعبیر مجموعه توزیع‌های احتمال سازگار با ساختار اعتقادی می‌توانند نقطه شروع مناسب‌تری برای توسعه روش های استنباط آماری به حوزه نظریه دمپستر–شفر باشد. به علاوه نشان داده شد که برای نظریه دمپستر–شفر به عنوان توسعه‌ای از احتمال می‌توان هر دو تعبیر ذهنی و عینی را ارائه کرد. مثال محاسبه امید ریاضی ۵۰ نشان داد که چگونه مفاهیم مفاهیم و ابزارهای معمول در نظریه احتمال قابل توسعه به نظریه دمپستر–شفر است.

از سوی دیگر، نشان داده شده است که اندازه امکان [۳۰] نیز یک اندازه از نوع موجه‌نمایی است که برای ساختاری اعتقادی هماهنگ [۳۱] با اعضای کانونی تو در تو [۳۲] حادث می‌شود [۸]. این یکی از دست‌آوردهای نظری اساسی در رابطه با نظریه دمپستر—شفر است. به این ترتیب نظریه دمپستر—شفر، در دو حالت خاص، نظریه‌های امکان واحتمال را در بر می‌گیرد.

گفتنی است که اسمتز [۳۳] چارچوب متفاوتی برای نظریه شواهد به نام مدل انتقال باور [۳۴] را معرفی می‌کند که در مسائل تصمیم‌گیری و استدلال و استنباط در قالب نظریه شواهد کاربردهای فراوانی دارد [۹]. بحث راجع به این چارچوب مهم مجالی جداگانه می‌طلبد و در این مختصر نمی‌گنجد.

در این مقاله فرض شده بود که مجموعه عالم سخن Ω مجموعه‌ای بسته است، به این معنی که تمامی حالت‌های ممکن تجربه تصادفی را در بر می‌گیرد. این فرض لزوما در تمامی مدل‌سازی‌های احتمالاتی معتبر نیست. مثال جالبی از عدم کارائی فرض جهان بسته [۳۵] خصوصا در کاربرد قاعده ترکیب دمپستر در [۱۰] آمده است. برای رفع این مشکل می‌توان به فرض جهان باز [۳۶] توسل جست که منجر به تخصیص جرم ناصفر به تهی می‌شود. مدل انتقال باور اسمتز راه حل مناسبی برای تصمیم گیری و استدلال و استنباط با فرض جهان باز به‌دست می‌دهد [٦].

توسعه‌های فازی متعددی برای نظریه دمپستر—شفر ارائه شده است. در تمامی این توسعه‌ها یک ساختار اعتقادی با اعضای کانونی فازی مد نظر بوده است. در این زمینه ین [۳۷] [۱۱] موفق به حفظ تعبیر حدود بالا و پایین احتمال با توسعه تعبیر رابطه سازگاری (۲) به یک رابطه سازگاری فازی شده است. به علاوه لوکس و اعرابی [۷] در توسعه فازی نظریه دمپستر—شفر به تعریفی با مقدار فازی [۳۸] برای توابع باور و موجه‌نمایی رسیده‌اند. در این توسعه مجموعه توزیع‌های احتمال سازگار با ساختار اعتقادی فازی به عنوان یک مجموعه فازی معرفی می‌شود که ویژگی مهم آن هم‌ارزی با توابع باور و موجه‌نمایی با مقدار فازی و نیز تابع جرم با اعضای کانونی فازی است، خاصیتی که در نظریه دمپستر—شفر وجود دارد اما در سایر توسعه‌های فازی آن دیده نمی‌شود.

نظریه حدود پایین و بالا احتمال دمپستر در فرم اولیه خود ناظر به فضاهای پیوسته نیز بوده است. اما محاسبه توابع باور و موجه‌نمایی در این حالت با پیچیدگی‌هایی همراه است که در حوصله این مقاله نیست. خواننده علاقه‌مند برای دیدن برخی ملاحظات و راه‌کارهای مرتبط می‌تواند به [۱۲] مراجعه کند.

مراجع

[1] Dempster, A.P. (1967), Upper and lower probabilities induced by a multi-valued mapping, Annl. Math. Stat., 38: 325-339.

---

[۳۰] Possibility measure
[۳۱] Consonant body of evidence
[۳۲] Nested focal elements
[۳۳] P. Smets
[۳۴] Transferable Belief Model
[۳۵] Closed world assumption
[۳۶] Open world assumption
[۳۷] J. Yen
[۳۸] Fuzzy valued

[2] Klir, G.J., Folger, T.A. (1988), Fuzzy Sets, Uncertainty, and Information. New Jersey: Prentice Hall.

[3] Lindley, D.V. (1980), L.J. Savage - His work in probability and statistics, Annl. Stat., 8: 1-24.

[4] Shafer, G. (1976), A Mathematical Theory of Evidence. Princeton, N.J.:Princeton University Press.

[5] Guan, J., Bell, D.A. (1991), Evidence Theory and Its Applications. Amsterdam: North-Holland.

[6] Smets, P. (1990), The combination of evidence in the transferable belief model, IEEE Trans. Pattern Anal. Machine Intell., 12: 447-458.

[7] Lucas, C., Araabi, B.N. (1999), Fuzzy generalization of the Dempster-Shafer theory: A fuzzy valued measure, IEEE Trans. Fuzzy Syst., 7: 255.270.

[8] Dubois, D., Prade, H. (1990), Consonant approximations of belief functions,Int. J. Approx. Reasoning, 4: 419-449.

[9] Smets, P., Kennes, R. (1994), The transferable belief model, Artificial Intelligence, 66: 191.243.

[10] Zadeh, L.A. (1984), A mathematical theory of evidence: Book review, AI Mag., vol. 5: 81-83, 1984.

[11] Yen, J. (1990), Generalizing the Dempster-Shafer theory to fuzzy sets, IEEE Trans. Syst. Man Cybern., 20: 559-570.

[12] Yen, J. (1992), Computing generalized belief function for continuous fuzzy sets, Int. J. Approx. Reasoning, 6: 1-31.