فصل 7

احتمال، امکان و تئوری شواهد

در فصل سوم برای بیان مفهوم اندازه‌گیری میزان فازی‌بودن، اندازه‌گیری فازی را به طور اجمالی معرفی کردیم. در این فصل به توضیح دقیق‌تر آن پرداخته، با معرفی تئوریهای احتمالات و امکان، به معرفی اندازه‌های احتمال[1]، امکان[2] و ضرورت[3] می‌پردازیم و در مورد تئوری شواهد بحث خواهیم کرد.

یکی از مطالب مهم این فصل، مقداردهی به حالتهای مختلف ابهام است. برای توضیح این نکته به دو نوع مشخص از ابهام اشاره می‌کنیم. فرض کنید یک متخصص در ایستگاه هواشناسی برای تشخیص وضع هوا، با استفاده از مشاهدات جوی به پیش‌بینی وضع هوا می‌پردازد. بدون‌شک در عملکرد این انسان متخصص مواردی وجود دارد که می‌توان ابهام یا عدم‌آگاهی را در آنها مشاهده کرد.

مثلاً اگر از او در مورد وضعیت آسمان از لحاظ ابری‌بودن سوال کنیم، برای پاسخگویی مناسب مجبور است (همان‌گونه که قبلاً در فصل اول توضیح داده شد) علاوه بر جواب صفر یا یک، مقادیر حقیقی بین این دو عدد را نیز به کار گیرد.

این پاسخ، یک ابهام در بردارد که از آن به *عدم وضوح* یاد می‌کنیم. این مورد از ابهام در دیگر مشاهدات متخصص نیز که به صورت حسی و دریافتی حاصل می‌شود وجود دارد. اگر از او در مورد میزان رطوبت سوال کنیم، متخصص پس از مراجعه به دستگاه اندازه‌گیری مربوط به میزان رطوبت هوا به ما می‌گوید: میزان رطوبت هوا زیاد است یا برابر با 75 درصد است. حال اگر از متخصص دیگری بخواهیم تا در همین وضعیت، میزان رطوبت را از همین دستگاه یا دستگاه دیگری بخواند، به ما پاسخی متفاوت ولی نزدیک به پاسخ اول خواهد داد.

نوع دیگری از عدم آگاهی یا ابهام در این مورد مشاهده می‌شود که به آن *عدم دقت*[4] می‌گوییم و در مواردی که اصطلاحاً اندازه‌گیری وجود دارد و به یک کمیت می‌توان مقادیر عددی (75.4 درصد، ...) و غیرعددی (زیاد، ...) نسبت داد، رخ می‌دهد.

---

[1] Probability
[2] Possibility
[3] Necessity
[4] Imprecision

اگر از تعدادی متخصص بخواهیم تا در شرایط یکسان با وسایل اندازه‌گیری خود میزان رطوبت هوا را گزارش کنند، پاسخهای متفاوتی را از آنها دریافت خواهیم کرد که با توجه به این پاسخهای دریافتی، می‌توان به هرکدام از اندازه‌های به‌دست‌آمده، با توجه به تکرار آن در بین پاسخها و همچنین تکرار اندازه‌های نزدیک به آن، یک بازه از درجه اعتبار را نسبت داد. یعنی یک حد پایین و یک حد بالا در فاصله $[0,1]$ برای اندازه در نظر بگیریم که به اولی *قابلیت اعتماد* و به دومی *قابلیت پذیرش* می‌گوییم .

به دست آوردن هر یک از این حدود از نوع اندازه‌گیری فازی بوده، حاصل آنها به عنوان اندازه‌های ضرورت و امکان شناخته می‌شوند. تفاوت این اندازه‌ها با اندازه‌های احتمال در ارتباط با خاصیت جمع‌پذیری می‌باشد. این اندازه‌ها خاصیت جمع‌پذیری ندارند، ولی به علت تشابه بازه $[0,1]$ این تصور به وجود می‌آید که این اندازه‌ها همان اندازه‌های احتمال می‌باشند. هرچندکه این تصور اشتباه است و μ_h , μ_l در ساختار کلی‌تر اندازه‌های فازی قرار می‌گیرند. اندازه‌های احتمال خود حالت خاصی از اندازه‌های فازی هستند.

در این فصل ضمن تعریف دقیق ریاضی نظریه امکان و احتمال و روشن‌تر شدن تفاوت میان آنها به تئوری شواهد خواهیم پرداخت. ابتدا نظریه شواهد را آنچنان که دمپستر معرفی کرد بیان می‌کنیم و سپس با ارائه دیدگاه شافر، خواهیم دید که به نتایج مشابهی می‌رسیم. می‌توان مدل دمپستر- شافر را بدون استفاده از هیچ مفهوم احتمالی به دست آورد. در پایان، قاعده ترکیب دمپستر را معرفی می‌کنیم و شیوه رسیدن به این قاعده را در نظریه حدود بالا و پایین دمپستر و نظریه شواهد شافر مورد بررسی قرار می‌دهیم.

7-1: تئوری امکان

اساس تئوری امکان بر ابهام و عدم دقت است که معمولاً در زبانهای طبیعی به طور ذاتی وجود دارد و به مفهوم امکان بیشتر از مفهوم احتمال نزدیک است. در مورد تفاوت موجود میان امکان و احتمال در مقدمه این فصل توضیحاتی داده شد. اکنون به تعریف ریاضی برای تئوری امکان می‌پردازیم.

به عنوان مثال، عبارت " X یک $\tilde{F}$ است" را در نظر بگیرید. در اینجا X نام یک شیء، متغیر یا یک عبارت است. مثلاً در عبارت " X یک عدد صحیح کوچک است"، X اسم یک متغیر است و در عبارت "سعید جوان است"، سعید نام یک موضوع یا شیء است.

$\tilde{F}$ یک مجموعه فازی (مثل عدد صحیح کوچک یا جوان) است که تابع عضویت آن با $\mu_{\tilde{F}}$ مشخص می‌شود. یک مفهوم اساسی در تئوری امکان، توزیع امکان می‌باشد که در نقطه مقابل توزیع احتمال قرار دارد و برای تعریف آن، ابتدا باید مفهوم تحدید فازی را تعریف کرد. برای داشتن تصوری از تحدید فازی، یک کیف با خاصیت الاستیک را در نظر بگیرید که با توجه به محتویاتش حجمهای ممکن را به‌اجبار می‌پذیرد، یک کیف با پوشش سخت را در نظر بگیرید که حجم آن یک عدد ثابت و صریح است و نیز یک چمدان نرم را تصور کنید که حجم محتویات آن، با توجه به نیرویی که محتویات چمدان به بدنه وارد می‌کند، تعیین می‌شود. در این مورد متغیر حجم کیف فرض می‌شود و مقادیری که این متغیر ممکن است داشته باشد $u \in U$ بوده، درجه‌ای که هر متغیر X می‌تواند به ازای مقادیر مختلف u بپذیرد، برابر با $\mu_{\tilde{F}}$ خواهد بود.

تعریف ۷-۱: اگر $\tilde{F}$ مجموعه فازی بر مجموعه مرجع U بوده، دارای تابع عضویت $\mu_{\tilde{F}}$ باشد، آنگاه $\tilde{F}$ یک *تحدید فازی*[1] بر متغیر X است اگر $\tilde{F}$ روی مقادیری که ممکن است X اختیار کند، به عنوان یک نیروی الاستیک عمل نماید. با فرض نسبت‌دادن مقدار u به متغیر X به صورت:

$$\mathrm{X} = u : \mu_{\tilde{F}}(u)$$

$\mu_{\tilde{F}}(u)$ بیانگر درجه فشار یا نیرویی است که وقتی مقدار u به X نسبت داده شد، توسط $\tilde{F}$ اعمال شده است.

اگر $\tilde{R}(X)$ یک تحدید فازی اعمال‌شده بر X باشد، به صورت $\tilde{R}(X) = \tilde{F}$ بیان خواهد شد که مجموعه فازی $\tilde{F}$ را به تحدید فازی $\tilde{R}(X)$ نسبت می‌دهد.

[1] Fuzzy restriction

فرض کنید که $A(X)$ یک نشان یا خاصیت از متغیر X باشد. مثلاً اگر داشته باشیم "سن سعید" $A(X)=$ و $\tilde{F}$ نیز یک مجموعه فازی مانند "جوان" باشد، عبارت "سعید جوان است" یا به تعبیر بهتر"سن سعید جوان (کم) است"، با توجه به تعریف فوق می‌تواند به صورت زیر بیان شود:

$$\tilde{R}(A(X))=\tilde{F}$$

مثال: اگر P عبارت "سعید جوان است" باشد و "جوان" یک مجموعه فازی بر مجموعه مرجع $U=[0,100]$ باشد که با تابع عضویت زیر مشخص می‌شود:

$$\mu_{young}(u)=S(u,20,30,40)$$

که u بیانگر متغیر سن بوده، تابع S نیز به صورت زیر تعریف شده است:

$$S(u,\alpha,\beta,\gamma)=\begin{cases}1 & if \quad u\le\alpha\\ 1-2\left(\dfrac{u-\alpha}{\gamma-\alpha}\right)^2 & if \quad \alpha\le u\le\beta\\ 2\left(\dfrac{u-\gamma}{\gamma-\alpha}\right)^2 & if \quad \beta\le u\le\gamma\\ 0 & if \quad u\ge\gamma\end{cases}$$

در این‌مورد $A(X)$ برابر با (سعید) Age می‌باشد. یعنی عبارت "سعید جوان است" به صورت زیر بیان می‌شود:

$\tilde{R}$ (Age (سعید)) = جوان

فرض کنید سن سعید 19 سال باشد یعنی u = 19 بوده، درجه عضویت سن 19 سال در مجموعه فازی "جوان" برابر با 0.7 باشد. با توجه به تعریف فوق، این درجه عضویت منتقل می‌شود و درجه امکان عبارت "سعید جوان است" برابر با 0.7 خواهد شد.

حال با روشن‌شدن مفهوم تحدید فازی، به تعریف توزیع امکان می‌پردازیم:

تعریف 7-2: اگر $\tilde{F}$ یک مجموعه فازی بر مرجع U بوده که با تابع عضویت $\mu_{\tilde{F}}(u)$ مشخص شده است و X یک متغیر باشد که می‌تواند مقادیر متعلق به مجموعه مرجع U را دریافت کند، همچنین $\tilde{F}$ یک تحدید فازی $\tilde{R}(X)$ باشد که بر متغیر X اعمال شده است

آنگاه "متغیر X یک $\tilde{F}$ است" که به صورت $\tilde{R}(X)=\tilde{F}$ بیان می‌شود، یک توزیع امکان را تعریف می‌کند که برابر با $\tilde{R}(X)$ فرض می‌شود.

تابع توزیع امکان $\prod_X(u)$ که توزیع امکان $\prod_X$ را مشخص می‌نماید به صورت عددی برابر با تابع عضویت $\mu_{\tilde{F}}(u)$ تعریف می‌شود:

$$\prod\nolimits_X = \mu_{\tilde{F}}$$

مثال: اگر مجموعه مرجع U بیانگر اعداد صحیح مثبت و $\tilde{F}$ مجموعه فازی اعداد صحیح کوچک با تعریف زیر باشد:

$$\tilde{F}=\{(1,1),(2,1),(3,0.8),(4,0.6),(5,0.4),(6,0.2)\}$$

آنگاه عبارت " X یک عدد صحیح کوچک است» به صورت توزیع امکان بر X با تعریف زیر بیان می‌شود:

$$\prod\nolimits_X = \tilde{F}$$

مثلا $(3,0.8)$ نشان می‌دهد که اگر X مقدار 3 را بپذیرد، درجه امکان " X یک عدد کوچک است" برابر با 0.8 می‌باشد.

برای روشن‌تر شدن مفهوم توزیع امکان، مثال مطرح‌شده توسط زاده (1978) را با اندکی تغییر بررسی می‌نماییم. در این مثال، مفهوم توزیع احتمال و توزیع امکان به خوبی از یکدیگر تفکیک شده‌اند.

مثال: عبارت "زهرا X تخم‌مرغ برای صبحانه می‌خورد" را با $X=\{1,2,...\}$ در نظر می‌گیریم. برای مجموعه X یک توزیع امکان و نیز یک توزیع احتمال قابل بیان می‌باشد. توزیع امکان $\prod_X(u)$ می‌تواند با درجه سادگی خورده‌شدن u تخم مرغ توسط زهرا بیان شود ولی برای به دست آوردن توزیع احتمال $P_x(u)$ باید از روش دیگری استفاده کنیم.

مثلاً میزان تخم‌مرغهای مصرف‌شده در 100 صبحانه زهرا را یادداشت کرده تا با توجه به تعداد تخم‌مرغهای مصرف‌شده در هر دفعه، احتمال $P_x(u)$ را پیش‌بینی نماییم. می‌توان مقادیر $\prod_X(u)$ و $P_x(u)$ را به صورت زیر در نظر گرفت:

u	1	2	3	4	5	6	7	8

$\prod_X(u)$	1	1	1	1	0.8	0.6	0.4	0.2
$P_x(u)$	0.1	0.8	0.1	0	0	0	0	0

مشاهده می‌کنیم که اندازه‌های امکان، یک حد بالا برای اندازه‌های احتمال می‌باشند. در فصل سوم، اندازه‌گیری امکان را برای حالتی که A یک مجموعه کلاسیک و صریح باشد تعریف کردیم. حال با فرض اینکه $\tilde{A}$ یک مجموعه فازی است، تعریفی عمومی‌تر برای اندازه‌گیری امکان ارائه می‌دهیم:

تعریف 7-3: اگر $\tilde{A}$ یک مجموعه فازی بر مجموعه مرجع U باشد و یک توزیع امکان Π_X تعریف‌شده بر متغیر X موجود باشد که بتواند مقادیر متعلق به مجموعه مرجع U را دریافت نماید، اندازه گیری امکان یا $\prod_X(\tilde{A})$ به صورت زیر تعریف می‌شود:

$$poss\{\text{X} \quad is \quad \tilde{A}\} = \prod_X(\tilde{A})$$
$$= \sup_{u \in U} \; min\{\mu_{\tilde{A}}(u), \prod_X(u)\}$$

مثال: اگر توزیع امکان حاصل از عبارت " X یک عدد صحیح کوچک مثبت است" در مثال قبل را در نظر بگیریم، خواهیم داشت:

$$\Pi_X = \{(1,1),(2,1),(3,0.8),(4,0.6),(5,0.4),(6,0.2)\}$$

و اگر مجموعه کلاسیک $A = \{3,4,5\}$ مفروض باشد، اندازه‌گیری امکان $\prod(A)$ برابر خواهد بود با:

$$\prod(A) = max(0.8,0.6,0.4) = 0.8$$

اگر $\tilde{A}$ بیانگر مجموعه فازی "اعداد صحیحی که کوچک نیستند" باشد:

$$\tilde{A} = \{(3,0.2),(4,0.4),(2,1),(5,0.6),(6,0.8),(7,1)\}$$

اندازه‌گیری امکان" X یک عدد صحیح کوچک نیست" برابر خواهد بود با:

$$\text{Poss}\ (\text{ X is not a small integer}) = max\ \{0.2,0.4,0.4,0.2\} = 0.4$$

با توجه به تعریف 3-2 برای اندازه‌گیری امکان، که در آن Ω یک زیرمجموعه کاملاً امکان‌پذیر از مجموعه مرجع در نظر گرفته شده بود، خواهیم داشت:

$$g(0)=0 \quad , \quad g(\Omega)=1$$

و همین‌طور با توجه به شرط دوم تعریف 3-2 که عبارت بود از:

$$A \subseteq B \Rightarrow g(A) \le g(B)$$

خواهیم داشت:

$$g(A \cup B) \ge max\big(g(A), g(B)\big)$$
$$g(A \cap B) \le min\big(g(A), g(B)\big) \quad for \quad A,B \subseteq \Omega$$

اندازه‌گیری امکان با تعریف فوق در مواردی محدود به صورت زیر خواهد بود:

$$\prod(A \cup B) = max\Big(\prod(A), \prod(B)\Big)$$
$$\prod(A \cap B) = min\Big(\prod(A), \prod(B)\Big)$$

که اگر $\not\subset A$ مکمل A در Ω در نظر گرفته شود، خواهیم داشت:

$$\prod(A \cup \not\subset A) = \max\Big(\prod(A), \prod(\not\subset A)\Big) = 1$$

که بیانگر این است که یا مجموعه A یا مکمل آن، کاملاً امکان‌پذیر می‌باشند. در تئوری امکان یک اندازه‌گیری دیگر تعریف می‌شود که دوگان اندازه‌گیری امکان می‌باشد:

$$N(A \cap B) = min\big(N(A), N(B)\big)$$

N *اندازه گیری ضرورت* نامیده می‌شود و $N(A)=1$ نشان‌دهنده آن است که درستی A ضروری است و A کاملاً قابل اعتماد می‌باشد. رابطه دوگانی میان اندازه‌گیری امکان و ضرورت به صورت زیر می‌باشد:

$$\prod(A) = 1 - N(\not\subset A) \qquad \forall A \subseteq \Omega$$

اندازه‌های ضرورت شرط زیر را برآورده می‌کنند:

$$min\big(N(A), N(\not\subset A)\big) = 0$$

و رابطه بین اندازه‌های ضرورت و اندازه‌های امکان طبق شرایط زیر است:

$$\prod(A) \geq N(A) \qquad \forall A \subseteq \Omega$$
$$N(A) \geq 0 \Rightarrow \prod(A) = 1$$
$$\prod(A) < 1 \Rightarrow N(A) = 0$$

در روابط فوق Ω همیشه یک مجموعه متناهی فرض شده است.

مثال: فرض کنید در یک کلاس ۶ دانش آموز داریم که با توجه به امتحانات گذشته جدول زیر را به عنوان امکان نمرات A تا E برای آنان در نظر گرفته‌ایم:

نمرات

Student	A	B	C	D	E
1	0.8	1	0.7	0	0
2	1	0.8	0.6	0.1	0
3	0.6	0.7	0.9	1	0
4	0	0.8	0.9	0.5	0
5	0	0	0.3	1	0.2
6	0.3	1	0.3	0	0

قبل از هر چیز باید توجه داشته باشیم که تابع عضویت نمرات دانش‌آموز شماره چهار یک تابع ممکن نیست، زیرا $g(\Omega) \neq 1$است.

حال سؤالات زیر را مطرح می‌نماییم:

1- چقدر می‌توان اطمینان داشت که دانش‌آموز شماره یک در امتحان بعدی نمره B کسب کند؟

در این مورد "A"، مجموعه $\{B\}$ بوده و "$\not\subset A$" ، $\{A,C,D,E\}$ می‌باشد بنابراین:

$$pls(\{B\}) = \prod(\{B\}) = 1$$
$$bel(\{B\}) = N(\{B\}) = min\{1 - \prod_i\} = min\{0.2, 0.3, 1, 1\} = 0.2$$

بنابراین امکان آنکه دانش‌آموز شماره یک نمره B کسب کند برابر با 1 بوده، ضرورت آن برابر با 0.2 خواهد بود.

2- اگر بخواهیم میزان درستی عبارت" یکی از دو دانش‌آموز شماره یک و دو نمره A یا B کسب کند" را به دست آوریم، باید Ω را به صورت متفاوتی تعریف کنیم و اعضای آن را عناصر دو سطر اول جدول در نظر بگیریم که نتیجه به صورت زیر خواهد شد:

$$pls(A)=\prod(A)=\prod(student\ 1\ A\ \ or\ \ B\ student\ 2\ A\ \ or\ \ B)=1$$

$$bell(A)=N(A)=min\{0.3,1,1,0.4,0.9,1\}=0.3$$

با توجه به معرفی اندازه‌گیری امکان و ضرورت، در مورد آنها خواهیم داشت:

$$N(A)+N(\not\subset A)\leq 1$$

$$\prod(A)+\prod(\not\subset A)\geq 1$$

درحالی‌که در اندازه گیری احتمال کلاسیک داریم:

$$p(A)+p(\not\subset A)=1$$

و با شرط $A\cap B=\varnothing$ خواهیم داشت:

$$p(A\cup B)=p(A)+p(B)$$

در بخش تئوری شواهد در این مورد بیشتر توضیح داده، نشان خواهیم داد:

$$N(A)\leq p(A)\leq \prod(A)\qquad \forall A\subseteq\Omega$$

یعنی اندازه‌های امکان و ضرورت محدودکننده‌های اندازه احتمال می‌باشند.

2-7: احتمال پیشامدهای فازی

با توضیحات ارائه شده در قسمتهای قبل می‌دانیم امکان، جانشین و جایگزین احتمال نیست بلکه نوع دیگری از ابهام است. اگر فرض کنیم که یک پیشامد نیز به صورت صریح و قطعی تعریف نشده باشد بلکه به صورت یک توزیع امکان یا یک مجموعه فازی تعریف شده باشد، در این صورت با یک ابهام روبرو هستیم. زیرا این پیشامد که مبهم تعریف شده است به صورت قطعی اتفاق نخواهد افتاد و مفهوم احتمال برای آن نیز قابل تعریف است. اما این مفهوم از احتمال را به دو صورت می‌توان بیان کرد. در حالت اول، احتمال وقوع یک پیشامد

فازی، یک عدد قطعی است و برای حالت دیگر، می‌توان احتمال وقوع رویداد را نیز یک مجموعه فازی در نظر گرفت.

7-2-1: احتمال یک پیشامد فازی به عنوان عدد اسکالر

در تئوری کلاسیک احتمالات، پیشامد A یک عضو از مجموعه a با میدان α می‌باشد که از زیرمجموعه‌های فضای نمونه Ω انتخاب شده است. یک اندازه احتمال P یک اندازه نرمال‌شده بر فضای قابل اندازه‌گیری (Ω, a) می‌باشد که در آن P، یک تابع با مقدار حقیقی است که به هر پیشامد A یک اندازه احتمال $P(A)$ با شرایط زیر را نسبت می‌دهد:

۱. $A \in a \qquad p(A) \geq 0$

۲. $p(\Omega) = 1$

۳. اگر $A_i \in a$, $i \in I \subset \mathbb{N}$ باشد، به‌گونه‌ای که A_i ها دوبه‌دو جدا از هم باشند داریم:

$$P\left(\bigcup_{i\in I} A_i\right) = \sum_{i\in I} P(A_i)$$

به عنوان مثال اگر Ω یک فضای n بعدی اقلیدسی و a یک زیرمجموعه از مجموعه‌ها با میدان α در $\mathbb{R}^n$ باشد، آنگاه احتمال پیشامد A می‌تواند به صورت زیر بیان شود:

$$P(A) = \int_A dp$$

اگر $\mu_A(x)$ تابع مشخصه مجموعه صریح A بوده، همچنین $E_p(\mu_A)$ امید ریاضی $\mu_A(x)$ باشد، آنگاه

$$P(A) = \int_{R^n} \mu_A(x)\,dp = E_p(\mu_A)$$

حال اگر فرض کنیم μ_A به جای تابع مشخصه، تابع عضویت یک مجموعه فازی باشد، قاعدتاً تعریف احتمال پیشامد A نباید تغییر کند. زاده (1968) احتمال پیشامد فازی $\tilde{A}$ با تابع عضویت $\mu_{\tilde{A}}(x)$ را به صورت زیر تعریف کرد:

تعریف 7-4: اگر $(\mathbb{R}^n, a, p)$ بیانگر یک فضای احتمالی باشد که در آن a شامل زیرمجموعه‌ای از زیرمجموعه‌ها با میدان α در $\mathbb{R}^n$ بوده، P یک اندازه احتمال بر $\mathbb{R}^n$ باشد، آنگاه یک پیشامد فازی در $\mathbb{R}^n$ یک مجموعه فازی $\tilde{A}$ در $\mathbb{R}^n$ خواهد بود که تابع عضویت آن را با $\mu_{\tilde{A}}(x)$ نمایش می‌دهیم. می‌توان احتمال پیشامد فازی $\tilde{A}$ را به وسیله انتگرال لبسک[1]- استیلتجز[2] به صورت زیر بیان کرد:

$$P\left(\tilde{A}\right) = \int_{R^n} \mu_{\tilde{A}}\left(x\right) dp = E_p\left(\mu_{\tilde{A}}\right)$$

زاده شباهتهای بین احتمال پیشامدهای فازی و احتمال پیشامدهای قطعی را برشمرد. در آن زمان ایشان نظریه خود را بسیار قابل‌پذیرش می‌دانست ولی آن را به عنوان یک اصل موضوعه قبول نکرد. در سال 1982 اسمتز[3] نشان داد که در محیطهای غیرفازی، احتمال غیرفازی پیشامدهای فازی را می‌توان به عنوان یک اصل موضوعه قبول کرد. باید توجه داشت که علاوه بر تعریف فوق، تعاریف دیگری نیز برای احتمال یک رویداد فازی به عنوان عدد اسکالر ارائه شده است که در این کتاب از معرفی آنها خودداری کرده، به تعریف دیگری از احتمال پیشامد فازی می‌پردازیم.

7-2-2: احتمال یک پیشامد فازی به عنوان مجموعه فازی

در این قسمت فرض می‌کنیم همه مجموعه‌ها به تعداد متناهی عضو دارند و همچنین هر عضو $x_i \in X$ (X مجموعه مرجع است) دارای اندازه احتمال $P(x_i)$ می‌باشد.

$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) \mid x \in X\}$ یک مجموعه فازی است که به عنوان یک پیشامد فازی فرض شده است که $\mu_{\tilde{A}}(x)_i$ تابع عضویت $x_i \in \tilde{A}$ است. مجموعه‌های در سطح a یا برشهای a را نیز طبق تعریف فصل دوم می‌توان به صورت A_α نمایش داد.

---

[1] Lebesque
[2] Stieltjes
[3] Smets

یاگر در سال های 1979 و 1984 مطرح کرد که کاملاً طبیعی است که احتمال یک مجموعه در سطح a به صورت $P\left(A_\alpha\right)=\sum_{x\in A_\alpha} p\left(x\right)$ بیان شده باشد. او بر این اساس احتمال پیشامدهای فازی را این گونه تعریف کرد:

تعریف 7-5: اگر $\tilde{A}$ یک پیشامد فازی و A_α یک برش a از آن باشد، آنگاه احتمال پیشامد فازی $\tilde{A}$ می‌تواند به صورت زیر تعریف شود:

$$P_Y\left(\tilde{A}\right)=\left\{\left(P\left(A_\alpha\right),\alpha\right)\mid \alpha\in[0,1]\right\}$$

که می‌توان از آن به "حداقل احتمال پذیرش شرط $\tilde{A}$ با درجه a" تعبیر کرد. حرف Y در عبارت P_Y نشان‌دهنده آن است که P_Y اندازه احتمال ناشی از تعریف یاگر است تا با اندازه احتمال تعریف‌شده توسط زاده (با نماد P) اشتباه نشود.

یاگر یک تعریف دیگر نیز از احتمال پیشامدهای فازی ارائه داده است:

تعریف ۷-٦ : درستی عبارت "احتمال پیشامد $\tilde{A}$ حداقل برابر w است" به صورت مجموعه فازی $P_Y^*\left(\tilde{A}\right)$ با تابع عضویت زیر بیان می‌شود:

$$P_Y^*\left(\tilde{A}\right)\left(w\right)=\sup_\alpha\left\{\alpha\mid P\left(A_\alpha\right)\geq w\right\},w\in[0,1]$$

باید توجه داشت که در این حالت، مشخص‌کننده ارزش اندازه احتمال برخلاف تعریف قبل که a بود، مقدار w می‌باشد.

ضمناً نمایش تابع عضویت در این تعریف، براساس نمادهای مورد استفاده یاگر صورت گرفته که با $P_Y^*\left(\tilde{A}\right)\left(w\right)$ نشان داده شده است.

مکمل مجموعه $\tilde{A}$ را با $\not\subset\tilde{A}$ نمایش می‌دهیم: $\not\subset\tilde{A}=\left\{\left(x,1-\mu_{\tilde{A}}\left(x\right)\right)\mid x\in X\right\}$ و مجموعه‌های در سطح a برای $\not\subset\tilde{A}$ با $\left(\not\subset\tilde{A}\right)_\alpha$ بیان می‌شوند. آنگاه

$$P_Y^*\left(\not\subset\tilde{A}\right)\left(w\right)=\sup_\alpha\left\{\alpha\mid P\left(\not\subset\tilde{A}_\alpha\right)\geq w\right\},w\in[0,1]$$

می‌تواند به درستی عبارت " احتمال عدم اتفاق پیشامد $\tilde{A}$ حداقل برابر wاست" تعبیر شود. درستی "احتمال پیشامد $\tilde{A}$ حداکثر برابر w باشد" را می‌توان با $\bar{P}_Y^*\left(\tilde{A}\right)\left(w\right)$ نمایش داد که رابطه زیر در مورد آن برقرار است:

$$\bar{P}_Y^*\left(\tilde{A}\right)=1-P_Y^*\left(\not\subset \tilde{A}\right)$$

حال می‌توانیم بگوییم که ترکیب "و" برای دو عبارت " احتمال پیشامد $\tilde{A}$ حداقل برابر w است" و "احتمال پیشامد $\tilde{A}$ حداکثر برابر w است" می‌تواند به صورت عبارت "احتمال پیشامد $\tilde{A}$ دقیقاً برابر w است" بیان شود. اگر $\bar{P}_Y^*\left(\tilde{A}\right),P_Y^*\left(\tilde{A}\right)$ به صورت توزیعهای امکان در نظر گرفته شده باشند، در این صورت ترکیب عطفی آنها همان اشتراک دو مجموعه فازی است که می‌تواند با عملگر *min* به دست بیاید.

تعریف 7-7: اگر $\bar{P}_Y^*\left(\tilde{A}\right),P_Y^*\left(\tilde{A}\right)$ را با تعریف فوق در نظر گرفته باشیم، توزیع امکان بیان‌شده با عبارت "احتمال پیشامد $\tilde{A}$ دقیقاً برابر w است" می‌تواند به صورت زیر بیان شود:

$$\bar{P}_Y\left(\tilde{A}\right)\left(w\right)=min\left\{P_Y^*\left(\tilde{A}\right)\left(w\right),\bar{P}_Y^*\left(\tilde{A}\right)\left(w\right)\right\}$$

مثال: پیشامد فازی $\tilde{A}=\left\{\left(x_1,1\right),\left(x_2,0.7\right),\left(x_3,0.6\right),\left(x_4,0.2\right)\right\}$ با احتمال $p_4=0.2,p_3=0.3,p_2=0.4,p_1=0.1$ برای هر عضو آن تعریف شده است و نیز $p\left\{x_2\right\}=0.4$ است، درحالی‌که x_2 با درجه 0.7 به پیشامد $\tilde{A}$ تعلق دارد.

ابتدا با مشخص کردن برشهای α به عنوان A_α برای همه $\alpha\in\left[0,1\right]$ شروع می‌کنیم. در گام بعدی، احتمال پیشامدهای صریح A_α را محاسبه کرده، بازه‌هایی از w را برای $p\left(A_\alpha\right)\geq w$ در نظر می‌گیریم. نهایتاً $P_Y^*\left(\tilde{A}\right)$را با مقدار بالاتر α در بازه خود به دست می‌آوریم. محاسبات فوق در زیر به طور خلاصه به نمایش درآمده است:

α	A_α	$P\left(A\right)$	w	$P_Y^*=\sup\alpha$

$[0,.2]$	$\{x_1,x_2,x_3,x_4\}$	1	$[.8,1]$	.2
$[.2,.6]$	$\{x_1,x_2,x_3\}$	.8	$[.5,.8]$	.6
$[.6,.7]$	$\{x_1,x_2\}$	.5	$[.1,.5]$	.7
$[.7,1]$	$\{x_1\}$	.1	$[0,.1]$	1

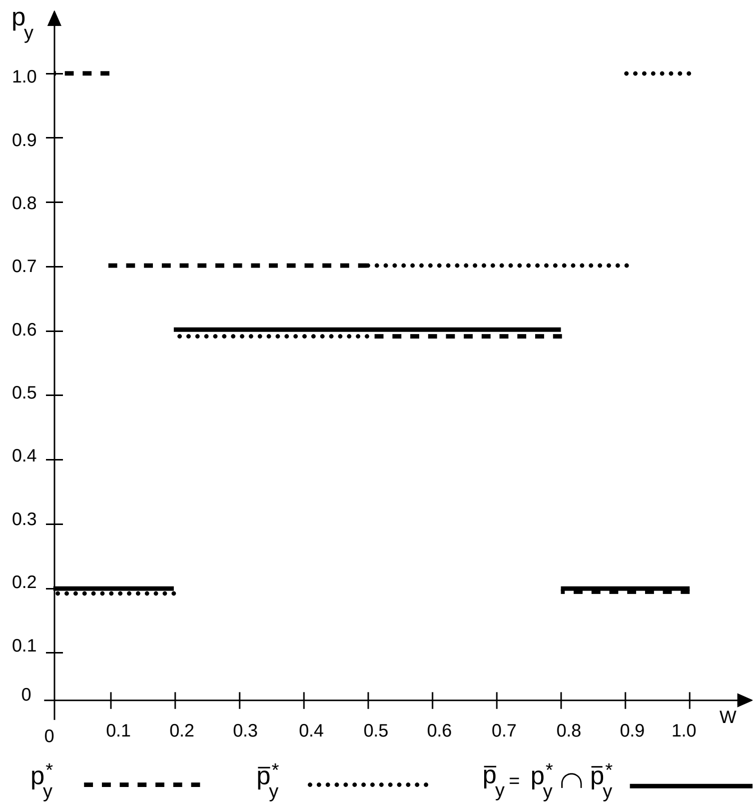
به صورت مشابه $\bar{P}_Y^*\left(\tilde{A}\right)=1-P_Y^*\left(\not\subset A\right)$ را به دست می‌آوریم:

	$\left(\not\subset \tilde{A}\right)_\alpha$	$P\left(\not\subset \tilde{A}\right)_\alpha$	w	$\bar{P}_Y^*\left(\not\subset \tilde{A}\right)$	$\bar{P}_Y^*\left(\tilde{A}\right)=1-P_Y^*\left(\not\subset \tilde{A}\right)$
0	$\{x_1,x_2,x_3,x_4\}$	1	$[0.9,1]$	0	.1
$[0,.3]$	$\{x_2,x_3,x_4\}$	.9	$[.5,.9]$	.3	.7
$[.3,.4]$	$\{x_3,x_4\}$	.5	$[.2,.5]$	.4	.6
$[.4,.8]$	$\{x_4\}$	.2	$[0,.2]$	.8	.2
$[.8,1]$	0	0	0	1	0

احتمال $\bar{P}_Y\left(\tilde{A}\right)$ را با اشتراک مجموعه‌های فازی $\bar{P}_Y^*\left(\tilde{A}\right),\bar{P}_Y^*\left(\tilde{A}\right)$ (که با عملگر *min* مدل شده‌اند) می‌توان به دست آورد:

$$\bar{P}_Y\left(\tilde{A}\right)(w)=\begin{cases}0 & if \quad w=0\\ 0.2 & if \quad w=[0,0.2]\\ 0.6 & if \quad w=[0.2,0.8]\\ 0.2 & if \quad w=[0.8,1]\end{cases}$$

شکل 7-1 بیانگر مجموعه‌های فازی $\bar{P}_Y\left(\tilde{A}\right)(w),\bar{P}_Y^*\left(\tilde{A}\right),P_Y^*\left(\tilde{A}\right)(w)$ است.



شکل7-1: مجموعه‌های فازی $\bar{P}_Y\left(\tilde{A}\right)(w), \bar{P}_Y^*\left(\tilde{A}\right), P_Y^*\left(\tilde{A}\right)(w)$

3-7: تئوری شواهد

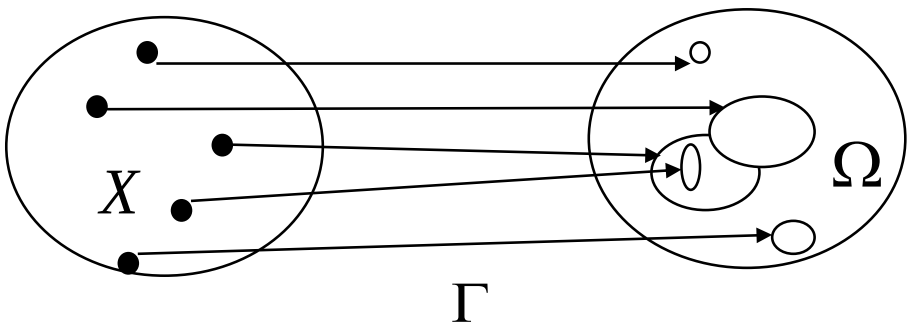
دمپستر و سویج[1] در مباحثات میان خود درگیر این مسئله مهم بوده‌اند که ساختار کلاسیک احتمال، امکان نمایش عدم آگاهی را ندارد، در صورتی که در بیان اطلاعات مسائل دنیای پیرامون، اکثر اوقات با عدم آگاهی در مورد بعضی از جنبه‌های آن (مثال ایستگاه هواشناسی در ابتدای همین فصل) مواجه هستیم.

---

[1] L.J.Savage

ضمناً احتمال یک اندازه جمع‌پذیر است، درحالی‌که مطالعات تجربی چنین خاصیتی را در مورد احتمال پدیده‌ها نشان نمی‌دهد و بیشتر به نظر می‌آید احتمال تجربی یک اندازه احتمال ترتیبی باشد. نظریه دمپستر[1] این مسأله را حل کرد. او دو اندازه احتمال ترتیبی و غیرجمع‌پذیر ارائه داد که مدل او امکان حمل عدم آگاهی را نیز دارا بود.

دمپستر در اواخر دهه 60 میلادی دیدگاه جدیدی را در بیان اندازه احتمال یک فضای Ω ارائه داد. او فرض کرد که اندازه احتمال روی فضای X موجود باشد و به علاوه، اعضای X به وسیله یک نگاشت چند مقداری Γ (مطابق شکل 7-2) روی زیرمجموعه‌های Ω تصویر شوند. حال سؤال این است که در مورد احتمال یک پیشامد روی Ω چه قضاوتی می‌توان داشت؟



شکل7-2: عدم آگاهی در نظریه دمپستر

دمپستر نشان داد که تعبیر احتمال یک پیشامد روی Ω مرتبط با نوعی عدم آگاهی است، به این ترتیب که به جای آنکه یک عدد به عنوان احتمال پیشامد مطرح شود، بازه‌ای از اعداد وجود دارد که می‌توانند احتمال پیشامد باشند. این بازه توسط یک حد بالا و پایین محدود می‌شود، در واقع دو اندازه μ_h, μ_l را روی Ω به دست می‌آوریم که برای آنها تعبیر حدود بالا و پایین احتمال را داریم. باید دقت داشت که این اندازه‌ها، اندازه احتمال به معنی کلاسیک آن نیستند، زیرا دارای دو خاصیت جمع‌پذیری و نرمال‌بودن نمی‌باشند.

[1] Dempster

در مورد اندازه‌های μ_h, μ_l داریم:

$$\mu_l(T)+\mu_l(R)\le \mu_l(T\cup R)+\mu_l(T\cap R)$$
$$\mu_l(T)+\mu_l(\bar{T})\le 1$$
$$\mu_h(T)+\mu_h(R)\le \mu_h(T\cup R)+\mu_h(T\cap R)$$
$$\mu_h(T)+\mu_h(\bar{T})\le 1$$

هر دو اندازه μ_h, μ_l دارای خاصیت یکنوایی می‌باشند:

$$\varnothing\subset T\subset R\subset\Omega\Rightarrow \mu_l(T)\le \mu_l(R)$$
$$\varnothing\subset T\subset R\subset\Omega\Rightarrow \mu_h(T)\le \mu_h(R)$$

و همچنین داریم:

$$\mu_l(\varnothing)=\mu_h(\varnothing)=0$$
$$\mu_l(\Omega)=\mu_h(\Omega)=1$$

یعنی می‌توان μ_h, μ_l را در ساختار کلی‌تر اندازه‌های فازی که در مورد آنها توضیح داده شد، در نظر گرفت. همان‌گونه که گفته شد این دو اندازه احتمال، ترتیبی و جمع‌پذیر هستند و امکان حمل عدم آگاهی را دارا می‌باشند.

بعدها در نیمه دوم دهه 70 میلادی، شافر[1] تئوری شواهد خود را مطرح کرد که نتیجه توسعه طبیعی آمار بیزی[2] بود و نتایج آن با نظریه حدود بالا و پایین دمپستر تطابق داشت.

تئوری شواهد در دهه 80 با توجه به کشف ارتباط میان آن و تئوری امکان و همچنین استفاده از آن به عنوان ابزاری مناسب برای بیان عدم قطعیت در سیستمهای خبره، بیش از پیش مورد توجه متخصصین علوم کامپیوتر قرار گرفت.

1-3-7: تئوری حدود بالا و پایین

فضای X را همراه با اندازه احتمال μ در نظر می‌گیریم که برای هر زیرمجموعه اندازه‌پذیر (متناهی) X مثل A، عدد $\mu(A)$ را به عنوان احتمال پیشامد نسبت می‌دهد. فضای Ω از طریق نگاشت چندمقداری Γ با X مربوط است و در آن داریم:

---

[1] Shafer

[2] Bayesian

$$\Gamma : X \to P_{\Omega} \quad ; \quad x \to \Gamma_x \subset \Omega$$

می‌خواهیم احتمال پیشامد $T \subset \Omega$ را بیابیم. به این منظور دو مجموعه $T^*, T_* \subset X$ را به صورت زیر تعریف می‌کنیم:

$$T^* = \{x \mid x \in X, \Gamma_x \cap T \neq \varnothing\}$$
$$T_* = \{x \mid x \in X, \Gamma_x \neq \varnothing, \Gamma_x \subset T\}$$

که در حالت خاص $T = \Omega$ داریم:

$$\Omega^* = \Omega_* = \{x \mid x \in X, \Gamma_x \neq \varnothing\}$$

که همان دامنه نگاشت Ω است.

حال می‌توانیم اندازه‌های μ_h, μ_l را روی Ω به صورت زیر تعریف کنیم:

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

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

با توجه به تعریف T^*, T_* داریم:

$$\forall T, T_* \subset T^* \Rightarrow \mu_l(T) \leq \mu_h(T)$$

در واقع $\mu(\mathrm{T}^*)$ بیانگر بیشترین مقدار احتمالی است که اندازه احتمال μ تحت نگاشت Γ می‌تواند به T منتقل کند. به همین ترتیب، $\mu(\mathrm{T}_*)$ بیانگر کمترین مقدار احتمالی است که تحت Γ از μ به T منتقل می‌شود. برای وضوح بیشتر مطلب فرض کنید:

$$\Omega = \{w_1, w_2, \dots, w_n\}$$

برای هر عضو در برد Γ می‌توان مجموعه ای از اعضای X را یافت که به آن عضو، نظیر می‌شود:

$$\Gamma_x = A \subset \Omega \Rightarrow \Gamma^{-1}(A) = \{x \mid x \in X, \Gamma_x = A\}$$

برد Γ زیرمجموعه‌ای از P_{Ω} می‌باشد، پس حداکثر تعداد اعضای آن برابر $2^n - 1$ است. فرض می‌کنیم برد Γ مجموعه زیر باشد:

$$\{A_1, A_2, \dots, A_l\} \qquad \varnothing \neq A_j \subset \Omega$$

پس می‌توان X را به صورت زیر افراز کرد:

$$\left\{\Gamma^{-1}\left(A_1\right),\Gamma^{-1}\left(A_2\right),\ldots,\Gamma^{-1}\left(A_l\right),\Gamma^{-1}\left(\varnothing\right)\right\}$$

بیانگر مجموعه اعضایی از X است که در دامنه Γ قرار ندارند. حال اگر به هر یک از مجموعه‌ها مقدار احتمال آن را نسبت دهیم، خواهیم داشت:

$$\left\{\mu\left(\Gamma^{-1}\left(A_1\right)\right),\mu\left(\Gamma^{-1}\left(A_2\right)\right),\ldots,\mu\left(\Gamma^{-1}\left(A_l\right)\right),\mu\left(\Gamma^{-1}\left(\varnothing\right)\right)\right\}$$

مجموع احتمالات فوق قطعاً برابر با یک خواهد بود، زیرا مجموعه‌های فوق در یک افراز قرار دارند. اگر فرض شود:

$$m\left(A_j\right)=\frac{\mu\left(\Gamma^{-1}\left(A_j\right)\right)}{1-\mu\left(\Gamma^{-1}\left(\varnothing\right)\right)},\qquad j=1,2,\ldots,l$$

آنگاه خواهیم داشت:

$$\sum_{j=1}^{l}\mu\left(\Gamma^{-1}\left(A_j\right)\right)+\mu\left(\Gamma^{-1}\left(\varnothing\right)\right)=1$$

$$\Rightarrow\sum_{j=1}^{l}\left[\frac{\mu\left(\Gamma^{-1}\left(A_j\right)\right)}{1-\mu\left(\Gamma^{-1}\left(\varnothing\right)\right)}\right]=1$$

$$\Rightarrow\sum_{j=1}^{l}m\left(A_j\right)=1$$

واضح است که $\mu\left(\Omega^*\right)=1-\mu\left(\Gamma^{-1}\left(\varnothing\right)\right)$ و نیز می‌توان دید که:

$$T^*=\bigcup_{A_j\cap T\neq\varnothing}\Gamma^{-1}\left(A_j\right)$$

$$T_*=\bigcup_{A_j\subset T}\Gamma^{-1}\left(A_j\right)$$

در نتیجه خواهیم داشت:

$$\mu\left(T^*\right)=\sum_{A_j\cap T\neq\varnothing}\mu\left(\Gamma^{-1}\left(A_j\right)\right)$$

$$\mu\left(T_*\right)=\sum_{A_j\subset T}\mu\left(\Gamma^{-1}\left(A_j\right)\right)$$

و نهایتاً داریم:

$$\mu_h(T)=\frac{\mu(T^*)}{\mu(\Omega^*)}=\sum_{A_j\cap T\neq\varnothing}\frac{\mu(\Gamma^{-1}(A_j))}{\mu(\Omega^*)}=\sum_{A_j\cap T\neq\varnothing}m(A_j)$$

$$\mu_l(T)=\frac{\mu(T_*)}{\mu(\Omega_*)}=\sum_{A_j\subset T}\frac{\mu(\Gamma^{-1}(A_j))}{\mu(\Omega^*)}=\sum_{A_j\subset T}m(A_j)$$

برای محاسبه μ_h,μ_l می‌توان فرض کرد که:

$$\Omega^*=\mathrm{X}\Leftrightarrow\mu(\Omega^*)=1\Leftrightarrow\Gamma^{-1}=(\varnothing)=\varnothing\Leftrightarrow\mu(\Gamma^{-1}(\varnothing))=0$$

پس $m(A_j)$ را می‌توان به عنوان احتمال پیشامدهای به دست آمده از افراز X به وسیله Γ در نظر گرفت. البته $m(A_j)$ها روی Ω دارای معنای دقیق‌تری می‌باشند که در اینجا به مفاهیم معنی‌دار حدود بالا و پایین احتمالی اکتفا می‌کنیم. خاصیتهای این حدود را می‌توان به صورت زیر بیان کرد:

$$\mu_l(T)\le\mu_h(T)\qquad\forall T\subset\Omega$$

$$\begin{cases}\mu_l(T)+\mu_l(\bar{T})\le1\\ \mu_h(T)+\mu_h(\bar{T})\ge1\end{cases}\qquad\forall T\subset\Omega$$

$$\mu_l(\varnothing)=\mu_h(\varnothing)=0\quad,\quad\mu_l(\Omega)=\mu_h(\Omega)=1$$

$$\mu_l(T)+\mu_h(\bar{T})=\mu_h(T)+\mu_l(\bar{T})=1$$

$$\mu_l(T)+\mu_l(R)\le\mu_l(T\cup R)+\mu_l(T\cap R)\le\mu_l(T)+\mu_h(R)\qquad\forall T,R\subset\Omega$$

$$\mu_l(T)+\mu_h(R)\le\mu_h(T\cup R)+\mu_h(T\cap R)\le\mu_h(T)+\mu_h(R)\qquad\forall T,R\subset\Omega$$

توسعه نامساوی فوق برای بیش از دو مجموعه امکان‌پذیر است که در حالت کلی داریم:

$$A_1,A_2,\dots,A_k\subset\Omega$$

$$\mu_L(A_1\cup A_2\cup\dots\cup A_k)\ge\sum_i\mu_l(A_i)-\sum_{i<j}\mu_l(A_i\cap A_j)+\dots+(-1)^{k+1}\mu_l(A_1\cap A_2\cap\dots\cap A_k)$$

$$\mu_h(A_1\cap A_2\cap\dots\cap A_k)\ge\sum_i\mu_h(A_i)-\sum_{i<j}\mu_h(A_i\cup A_j)+\dots+(-1)^{k+1}\mu_h(A_1\cup A_2\cup\dots\cup A_k)$$

2-3-7: تئوری شواهد شافر

در بخش 7-1 توضیحاتی در مورد اندازه‌های ضرورت و امکان داده شد، اما توضیح دقیق‌تر و کامل‌تر به این بخش موکول گردید.

در بخش قبل نیز با نظریه دمپستر، حدود بالا و پایین آن و $m\left(A_j\right)$ آشنا شدیم. در این بخش، تئوری شواهد شافر را بررسی می‌کنیم و خواهیم دید که دقیقاً به همان مفاهیم اما با تعبیری متفاوت خواهیم رسید.

با توجه به اینکه تئوری شواهد شافر توسعه طبیعی آمار بیزی[1] است، ابتدا به اختصار اصول آمار بیزی را معرفی کرده، سپس به تئوری شواهد می‌پردازیم.

فرض کنیم $\Omega=\left\{w_1,w_2,...,w_n\right\}$ مجموعه مرجع باشد، تابع $d:\Omega\rightarrow[0,1]$ را یک تابع چگالی بیزی می‌نامیم اگر $\sum_{w\in\Omega} d\left(w\right)=1$ باشد. یک تابع $bay:P_\Omega\rightarrow[0,1]$ که در آن P_Ω مجموعه توانی Ω است را با شرایط زیر، تابع بیزی (احتمال بیزی) می‌گویند:

$$bay\left(\varnothing\right)=0 \qquad bay\left(\Omega\right)=1$$

$$bay\left(A\cup B\right)=bay\left(A\right)+bay\left(B\right) \quad ,A\bigcap B=\varnothing \qquad ,A,B\subset\Omega$$

به جای رابطه دوم می‌توان از هر یک از روابط زیر به صورت معادل استفاده کرد:

$$bay\left(A\right)=\sum_{x\in A} bay\left(\left\{x\right\}\right) \qquad A\subset\Omega$$

$$baya2...\left(A\cup B\right)=bay\left(A\right)+bay\left(B\right)-bay\left(A\cap B\right) \qquad ,A,B\subset\Omega$$

$$bay\left(A_1\cup A_2\cup...\cup A_k\right)=\sum_{\varnothing\neq I\subset\{1,2,...,k\}}\left(-1\right)^{|I|+1} bay\left(\bigcap_{i\in I}A_i\right) \qquad ,\forall A_i\subset\Omega$$

لازم به ذکر است تابع احتمال بیزی، غیرنزولی است یعنی:

$$\overset{x\in A}{A}\subset B\rightarrow bay\left(A\right)\leq bay\left(B\right) \qquad \forall A,B\in\Omega$$

حال به رابطه بین تابع چگالی بیزی و تابع احتمال بیزی می‌پردازیم:

قضیه ۷-۱: اگر d یک تابع چگالی بیزی باشد آنگاه:

$$bay\left(A\right)=\sum_{x\in A} d\left(x\right) \qquad \forall x\in\Omega$$

یک تابع احتمال بیزی است.

---
[1] Bayesian statistics

و برعکس، اگر bay یک تابع احتمال بیزی باشد آنگاه:

$$d(x) = bay(\{x\}) \qquad \forall x \in \Omega$$

یک تابع چگالی بیزی می‌باشد.

حال که در مورد آمار بیزی آشنایی مختصری به دست آوردیم، به تئوری شواهد شافر می‌پردازیم. ابتدا تابع جرم و تابع اعتماد را تعریف می‌کنیم:

تعریف 7-8: تابع $m : P_\Omega \to [0,1]$ را یک *تابع جرم*[1] گوییم اگر:

$$m(\varnothing) = 0$$
$$\sum_{A \subset \Omega} m(A) = 1$$

تابع جرم به $\forall A \subset \Omega$ یک عدد $m(A)$ نسبت می‌دهد. $m(A)$ را جرم A یا *احتمال پایه* A می‌گوییم و $\{m(A)\}_{A \subset \Omega}$ را یک *تخصیص احتمال پایه* می‌نامیم. به $\forall A \subset \Omega$ که برای آن $m(A) > 0$ برقرار باشد، *عنصر کانونی تابع جرم* m گفته می‌شود.

اجتماع تمامی عضوهای کانونی m را *هسته*[2] m می‌نامیم. به مجموعه اعضای کانونی همراه با جرم آنها معمولاً یک *ساختار اعتقادی*[3] گفته می‌شود که دانستن آن، در واقع معادل با دانستن m فوق می‌باشد.

تعریف 7-9: تابع $bel : P_\Omega \to [0,1]$ را یک *تابع اعتماد* می‌نامیم اگر:

$$bel(\varnothing) = 0$$
$$bel(\Omega) = 1$$
$$bel(A_1 \cup A_2 \cup \ldots \cup A_k) \geq \sum_{\varnothing \neq I \subset \{1,2,\ldots,k\}} (-1)^{|I|+1} bay\left(\bigcup_{i \in I} A_i\right) \quad , \forall A_1, A_2, \ldots, A_K \subset \Omega$$

لازم به ذکر است تابع اعتماد، غیرنزولی است و برای $\forall A, B \subset \Omega$ داریم:

$$A \subset B \to bel(A) \leq bel(B)$$

---

[1] Mass function
[2] Core
[3] Belief network

مشابه ارتباط میان تابع چگالی و تابع احتمال بیزی که در قضیه 7-1 بیان شد، ارتباط میان تابع اعتماد و تابع جرم برقرار است که در قضیه 7-2 بیان می‌شود.

قضیه 7-2: اگر m یک تابع جرم باشد آنگاه تابع bel با تعریف:

$$bel(A)=\sum_{X\subset A}(-1)^{|A-X|}\ bel(x) \quad \forall A\subset\Omega$$

یک تابع اعتماد است و برعکس، یعنی اگر bel یک تابع اعتماد باشد آنگاه:

$$m(A)=\sum_{X\subset A}(-1)^{|A-X|}\ bel(x) \quad \forall A\subset\Omega$$

یک تابع جرم است.

علاوه بر تابع اعتماد و تابع جرم که در تناظر یک‌به‌یک با یکدیگر قرار دارند، می‌توان توابع دیگری را هم معرفی کرد که همان میزان آگاهی را شامل شوند. از مهمترین این‌گونه توابع می‌توان به تابع قابلیت پذیرش اشاره کرد.

تعریف 7-10: تابع $pls : P_\Omega \to [0,1]$ را یک تابع *قابلیت پذیرش* می‌نامیم اگر:

$$pls(\varnothing)=0$$

$$pls(\Omega)=1$$

$$pls(A_1\cap A_2\cap\ldots\cap A_k)\le\sum_{\varnothing\ne I\subset\{1,2,\ldots,k\}}(-1)^{|I|+1}\ pls\left(\bigcup_{i\in I}A_i\right) \quad ,\forall A_1,A_2,\ldots,A_K\subset\Omega$$

با تکیه بر مطالبی که قبلاً در بخش اندازه‌گیری امکان، ضرورت و احتمال تحت عنوان توابع N,Π بحث شد، داریم که اگر bel یک تابع اعتماد باشد آنگاه:

$$pls(A)=1-bel(\bar{A})$$

یک تابع قابلیت پذیرش است و اگر m تابع جرم نظیر bel باشد آنگاه:

$$pls(A)=\sum_{X\cap A\ne\varnothing}m(X)$$

مشاهده می‌شود که pls,bel در مقایسه با نظریه حدود بالا و پایین دمپستر همان تعریف μ_h,μ_l را دارند. پس تعبیر حد پایین و بالای احتمال را برای آنها قائل می‌شویم.

مقایسه تعریفها و قضیه آمار بیزی با تعریفها و قضیه تئوری شواهد که به دنبال هم آمدند، تشابه کامل دو تئوری را نشان می‌دهد. برای آنکه نشان دهیم نظریه شواهد، توسعه آمار بیزی است، کافی است نشان دهیم در حالت خاص به آن منجر می‌شود. بدین منظور با سه قضیه زیر، این امر را دنبال می‌کنیم.

قضیه۷-۳: اگر d تابع چگالی بیزی باشد، آنگاه m تعریف‌شده به صورت زیر:

$$m(\{x\}) = d(x) \qquad \forall x \in \Omega$$

$$m(X) = 0 \qquad \forall X \in \Omega \quad ; \quad |X| > 1$$

یک تابع جرم است و برعکس، اگر m یک تابع جرم باشد که:

$$m(X) = 0 \qquad \forall X \in \Omega \quad ; \quad |X| > 1$$

آنگاه d با تعریف:

$$d(x) = m(\{x\}) \qquad \forall x \in \Omega$$

یک تابع چگالی بیزی است.

قضیه7-4: اگر bay یک تابع احتمال بیزی باشد، آنگاه bay یک تابع اعتماد است و m تابع جرم نظیر این تابع اعتماد می‌باشد و خواهیم داشت:

$$m(\{x\}) = bay(\{x\}) \qquad \forall x \in \Omega$$

$$m(X) = 0 \qquad \forall X \in \Omega \quad ; \quad |X| > 1$$

قضیه7-5: اگر bel یک تابع اعتماد بوده، تابع جرم نظیر آن به صورت زیر باشد:

$$m(\{x\}) = bell(\{x\}) \qquad \forall x \in \Omega$$

$$m(X) \qquad \forall X \in \Omega \quad ; \quad |X| > 1$$

در این صورت bel یک تابع احتمال بیزی است و تابع چگالی نظیر آن:

$$d(x) = m(\{x\}) = bel(\{x\}) \qquad \forall x \in \Omega$$

پس آمار بیزی حالت خاصی است که اعضای کانونی ساختار اعتقادی خود اعضای Ω (نه زیر مجموعه‌های آن) هستند.

جالب است توجه داشته باشیم که بیان اخیر از تئوری شواهد به لحاظ پایه نظری کاملاً از نظریه حدود بالا و پایین دمپستر مستقل است. در واقع، هیچ ارجاعی به حدی برای احتمال بودن تابع اعتماد نداده‌ایم و آن را در یک ساختار اصل موضوعی بیان کرده‌ایم.

تعبیر bel به عنوان μ_l و یا pls به عنوان μ_h از ساختار نظریه نتیجه نشده است بلکه همسویی نتایج ما را به این نکات رسانده است، یعنی دو خاستگاه مختلف به نتایج یکسانی رسیده‌اند. پس می‌توان برای این نتایج هر دو معنا را قائل بود، یعنی bel هم تابع اعتماد و هم حد پایین احتمال است و m هم تابع جرم است و هم چگالی احتمال پیشامدهای Γ حاصل از افراز X به وسیله نگاشت چندمقداری Γ است.

در قالب تعاریف و قضایا برای توابع اعتماد و قابلیت پذیرش خواصی مطرح شد. از دیگر خواص pls,bel می‌توان به موارد زیر اشاره کرد که قبلاً در مورد اندازه‌های N,Π بیان شده‌اند:

$$pls(A)+pls(\tilde{A})\geq 1 \qquad \forall A\subset\Omega$$
$$bel(A)+bel(\tilde{A})\leq 1 \qquad \forall A\subset\Omega$$
$$bel(A)\leq pls(A) \qquad \forall A\subset\Omega$$

حال می‌توانیم بازه $\left[bel(A),pls(A)\right]$ را به عنوان بازه اعتماد تعریف کنیم که طول این بازه معرف عدم آگاهی است:

$$ignorance(A)=pls(A)-bel(A)$$

اگر به تعبیر حدود بالا و پایین احتمال توجه کنیم، بازه اعتماد تعبیری از بازه مقادیر ممکن احتمال است و $ignorance(A)$ عدم قطعیت در تعریف احتمال را نشان می‌دهد.

قضیه ۷-٦: هرگاه bel یک تابع اعتماد باشد، شرط لازم و کافی برای آنکه bel یک تابع احتمال بیزی باشد آن است که:

$$bel(A)=pls(A) \qquad \forall A\subset\Omega$$

یعنی در حالت خاص آمار بیزی، فرض بر این است که هیچ نوع عدم آگاهی نداریم:

$$ignorance(A)=0 \qquad \forall A\subset\Omega$$

به این ترتیب از دو راه متفاوت و با ایده‌ها و خاستگاههای مختلف، اندازه‌های *pls,bel* را به دست آورده‌ایم. این دو اندازه همراه با احتمال (حالت خاص آنها) متعلق به خانواده کلی‌تر اندازه‌های فازی هستند که بارها در این کتاب توضیح داده شده‌اند.

در پایان این قسمت یک مثال از دنیای واقعی ارائه می‌دهیم:

مثال: بیماری که از درد کمر رنج می‌برد به پزشک مراجعه می‌کند. پزشک پس از مشاهده حالات بیمار تشخیص می‌دهد که ممکن است آسیب‌دیدگی عصب سیاتیک باعث این بیماری در بدن شخص شده باشد، البته می‌توان احتمال داد عامل دیگری این عوارض را در شخص ایجاد کرده باشد. با این اطلاعات، یکی از شرایط زیر عامل بیماری می‌باشد:

آسیب دیدگی عصب سیاتیک (S) یا عوامل دیگر (O)

با توجه به اینکه مجموعه مرجع ما دارای دو عضو می‌باشد، چهار عضو کانونی داریم که در جدول 7-1 به نمایش گذارده شده‌اند. با توجه به تجربیات پزشک، می‌توان $m(A_i)$ را در ستون اول قرار داد که مجموع آن برابر با یک می‌باشد.

جدول7-1: یک مثال دنیای واقعی برای تئوری شواهد

A_i	$m(A_i)$	$bel(A_i)$	$pls(A_i)$
$\varnothing$	0	0	0
O	0.4	0.4	0.8
S	0.2	0.2	0.6
$O \cup S$	0.4	1	1

حال می‌خواهیم در این مشاهدات، درجه اعتماد و قابلیت پذیرش را به دست آوریم:

$$bel(O) = \sum_{B \subset O} m(B) = m(O) = 0.4$$

$$bel(S) = \sum_{B \subset S} m(B) = m(S) = 0.2$$

$$bel(O \cup S) = m(O) + m(S) + m(O \cup S) = 0.4 + 0.2 + 0.4 = 1$$

همچنین برای تابع قابلیت پذیرش داریم:

$$
\begin{aligned}
pls(O) &= \sum_{B\cap O\neq\varnothing} m(B) \\
&= m(O\cap O) + m\big(O\cap(O\cap S)\big) \\
&= 0.4+0.4=0.8 \\
pls(S) &= m(S\cap S) + m\big(S\cap(O\cap S)\big) \\
&= 0.2+0.4=0.6 \\
pls(O\cup S) &= m\big((O\cup S)\cap S\big) + m\big((O\cup S)\cap S\big) + m\big((O\cup S)(O\cup S)\big) \\
&= 0.4+0.2+0.4=1
\end{aligned}
$$

با توجه به عملیات فوق، پزشک متخصص می‌تواند تشخیص دهد که با درجه 0.2 قابل اعتماد است که عامل بیماری شخص، آسیب‌دیدگی عصب سیاتیک باشد که درجه قابلیت پذیرش آن 0.6 است. روشن است که در این مورد، عدم آگاهی برابر با 0.4 می‌باشد.

7-3-3: اعمال شروط و قاعده ترکیب

در این قسمت ترکیب دمپستر و شافر را معرفی می‌کنیم. فرض کنید فضاهای X_i با اندازه‌های احتمال μ_i برای $i=1,2,\ldots,n$ را داشته باشیم. اگر این فضاها مجزا بوده، اندازه‌های احتمال مستقل از یکدیگر باشند و نظیر هر فضای X_i نگاشت $\Gamma_i : X_i \to P_\Omega$ را در نظر بگیریم، مشاهده کردیم که هر Γ_i نظیر μ_i یک ساختار حد بالا و پایین احتمال را روی Ω ایجاد می‌کند. قاعده ترکیب دمپستر و شافر به بررسی مجموع اطلاعات حاصل از تابع مجزای X_i و توابع احتمال مستقل می‌پردازد.

7-3-3-1: روش دمپستر

با توجه به نظریه حدود بالا و پایین دمپستر، فرض می‌کنیم شرط $R\subset\Omega$ را روی فضای Ω داریم. می‌خواهیم احتمال پیشامد $T\subset\Omega$ را به شرط R بیابیم. با تعبیر دمپستر اندازه احتمال $T\subset\Omega$ روی X تغییر نمی‌کند و نگاشت Γ هم ثابت است. تنها مجموعه $\Omega-R$ ناممکن و در نتیجه نامحتمل خواهد بود، پس کافی است تا برد Γ را محدود به زیرمجموعه‌های R نماییم.

$$\Gamma'_X = \left(\Gamma_X\right) \cap R \qquad \forall x \in X$$

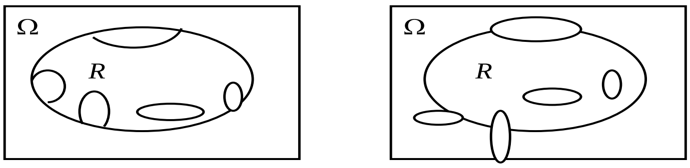
حال می‌توانیم حدود بالا و پایین احتمال را که توسط Γ'_X ایجاد می‌شوند، به عنوان حدود بالا و پایین احتمال شرطی روی فضای Ω به شرط R به صورت زیر تعریف کنیم:

$$\mu_h\left(T \mid R\right) = \frac{\mu_h\left(T \cap R\right)}{\mu_h\left(R\right)}$$

$$\mu_l\left(T \mid R\right) = 1 - \mu_h\left(\bar{T} \mid R\right) = \frac{\mu_l\left(T \cup \bar{R}\right) - \mu_l\left(\bar{R}\right)}{1 - \mu_l\left(\bar{R}\right)}$$

اگر برد Γ را $\{A_1, A_2, \ldots, A_l\}$ در نظر بگیریم و $m\left(A_j\right)$ های نظیر آن را داشته باشیم، برد Γ' برابر $\{A_1 \cap R, A_2 \cap R, \ldots, A_l \cap R\}$ بوده $m'\left(A_j \cap R\right)$ های نظیر آن مطابق فرمول زیر خواهد بود. مراحل اعمال شرط در شکل 7-3 قابل مشاهده است.

$$m'\left(A_j \cap R\right) = \begin{cases} \dfrac{m\left(A_j\right)}{1 - \sum\limits_{A_j \cap R \neq \varnothing} m\left(A_j\right)} & A_j \cap R \neq \varnothing \\ 0 & A_j \cap R \neq \varnothing \end{cases}$$



ب الف

ب: $A_j \cap R$ ها بعد از اعمال شرط الف: شرط R و A_j ها

شکل7-3: اعمال شرط

اگر دیدگاه اعمال شرط را گسترش دهیم، به قاعده ترکیب دمپستر می‌رسیم. یعنی برای فضاهای X_i با اندازه احتمال مستقل μ_i که $i = 1, 2, \ldots, n$ است و نظیر هر فضای X_i نگاشت Γ_i وجود دارد، با توجه به فرض استقلال آماری (اصل ضرب) خواهیم داشت:

$$X = X_1 \times X_2 \times ... \times X_n$$
$$\mu = \mu_1 \times \mu_2 \times ... \times \mu_n$$

و برای ترکیب نگاشتهای Γ_i با توجه به اینکه تصویر $x \in X$ روی Ω تنها شامل نقاطی است که تک‌تک اعضای $x = (x_1, x_2, ..., x_n)$ به آنجا تصویر شده‌اند، داریم:

$$\Gamma_x = \Gamma_{1x_1} \cap \Gamma_{2x_2} \cap ... \cap \Gamma_{nx_n} \qquad x = (x_1, x_2, ..., x_n) \in X$$

به بیان دیگر، اگر به هر نگاشت چندمقداری Γ_i به منزله یک تابع سازگاری C بین اعضای X_i و Ω با تعریف زیر نگاه کنیم:

$$C \subset X_i \times \Omega \qquad C(x_i, w) = \begin{cases} 1 & w \in \Gamma_i x_i \\ 0 & w \notin \Gamma_i x_i \end{cases}$$

زمانی که $C(x_i, w) = 1$ است، می‌گوییم x_i با w *سازگار* است و این واقعیت را با نماد $x_i C w$ نشان می‌دهیم. دراین صورت شرط سازگاری x با w آن است که تمام مؤلفه‌های x_i موجود در x با w سازگار باشند.

به این ترتیب، روش مناسبی برای ترکیب آگاهی‌های حاصل از n منبع مستقل، ارائه شده است. نکته اینجا است که با اضافه‌شدن هر آگاهی جدید، $|A_j|$ ها کوچک‌تر می‌شوند و تعداد اعضای موجود در عضو کانونی A_j کم می‌شود که این نکته با دیدگاه کلی ما در کاهش عدم‌قطعیت در مواجهه با اطلاعات جدید سازگار می‌باشد.

در هر صورت، این شیوه ترکیب را می‌توان برحسب $m(A_j)$ ها بیان کرد که به علت خاصیت شرکت‌پذیری مسأله ترکیب، دو منبع مستقل را مطرح می‌کنیم که البته برای منابع بیشتر (به علت وجود خاصیت شرکت‌پذیری) می‌توان عمل ترکیب را ادامه داد.

در مورد ترکیب دو منبع مستقل داریم:

$$\Gamma_1 : X_1 \to P_\Omega \qquad \mathrm{Im}\,age(\Gamma_1) = \{A_1, A_2, ..., A_l\}$$
$$\Gamma_2 : X_2 \to P_\Omega \qquad \mathrm{Im}\,age(\Gamma_2) = \{B_1, B_2, ..., B_k\}$$

با توجه به تعریف $m_2(B_j), m_1(A_i)$ داریم:

$$\mathrm{Im}\,age(\Gamma) = \{A_i \cap B_j \mid A_i \cap B_j \neq \varnothing; i = 1, 2, ..., k_l\}$$

و اگر $C \in \mathrm{Im}\,age(\Gamma)$ باشد، خواهیم داشت:

$$m_c = \frac{1}{N} \sum_{A_i \cap B_j = C} m_1(A_i) m_2(B_j)$$

که در آن:

$$N = \sum_{A_i \cap B_j \neq \varnothing} m_1(A_i) m_2(B_j) = 1 - \sum_{A_i \cap B_j \neq \varnothing} m_1(A_i) m_2(B_j)$$

Nضریب نرمال است تا مجدداً $\sum_{C \subset \Omega} m(C) = 1$ برقرار بشود.

اگر Γ_2 را به صورت $\Gamma_{2x_2} = \Omega \;\; ;\forall x_2 \in X_2$ تعریف کنیم برد Γ_2 تنها شامل یک عنصر یعنی Ω است و $m(\Omega) = 1$می‌باشد و $\Gamma_X = \Gamma_{1X_1}$ خواهد شد. به عبارت دیگر، این ترکیب هیچ تأثیری روی ساختار حدود بالا و پایین اولیه ندارد. با توجه به این امر، چنین منبعی را *منبع بدون اطلاعات* می‌نامیم. دقت شود که بدون اطلاعات بودن منبع به توزیع μ_2 درون آن ربطی ندارد و به نسبت نوع خاص نگاشت Γ_2 مربوط است.

به همین ترتیب، در یک حالت خاص دیگر اگر $\Gamma_{2x_2} = R \;\; ;\forall x_2 \in X_2$ فرض شود باز هم برد Γ_2 تنها شامل یک عنصر یعنی R است و $m(R) = 1$ می‌باشد و خواهیم داشت:

$$\Gamma_X = \Gamma_{1X_1} \cap R \qquad \forall x \in X$$

که این همان رابطه حدود بالا و پایین مشروط به R می‌باشد. یعنی در حالت خاص، ترکیب در آگاهی به همان نتیجه اعمال شرط R می‌رسد. دقت شود که قاعده اعمال شرط به عنوان حالت خاص قاعده ترکیب تعریف نشده است، بلکه چنین خاصیتی را بروز می‌دهد.

7-3-3-2: روش شافر

حال به معرفی روش شافر می‌پردازیم که اگرچه به طریق دیگری عمل می‌کند، ولی به همان نتایج دمپستر می‌رسد. دو ساختار اعتقادی را که با توابع جرم m_1 و m_2 معرفی شده‌اند در نظر می‌گیریم. جمع متعامد این دو ساختار مطابق تعریف زیر بیان می‌شود:

تعریف 7-11: توابع جرم m_1 و m_2 را روی Ω در نظر می‌گیریم:

$$m_1 : P_\Omega \to [0,1] \qquad m_1(\varnothing)=0 \qquad \sum_{A\subset\Omega} m_1(A)=1$$
$$m_2 : P_\Omega \to [0,1] \qquad m_2(\varnothing)=0 \qquad \sum_{B\subset\Omega} m_2(B)=1$$

جمع *متعامد* این دو به صورت زیر تعریف می‌شود:

$$(m_1 \oplus m_2)(c) = \frac{1}{N} \sum_{A_i \cap B_j = C} m_1(A) m_2(B)$$

که در آن:

$$N = 1 - \sum_{A\cap B\neq\varnothing} m_1(A) m_2(B)$$

اگر جمع متعامد m_1 و m_2 جرم را $m = m_1 \oplus m_2$ بنامیم آنگاه m خود یک تابع جرم روی Ω خواهد بود که m ترکیب دو ساختار اعتقادی با توابع جرم m_1 و m_2 است. یعنی ساختار اعتقادی حاصل از ترکیب، دارای تابع جرم m است که از جمع متعامد m_1و m_2 به دست آمده است. حال باید ترکیب‌پذیری دو ساختار اعتقادی مورد بحث قرار گیرد. دو ساختار اعتقادی که دارای شرط $\sum_{A\cap B\neq\varnothing} m_1(A) m_2(B) \neq 1$ باشند را ترکیب‌پذیر گوییم زیرا در غیر این صورت $N=0$ شده، جمع متعامد m_1 و m_2 قابل‌تعریف نخواهد بود.

تعریف 7-12: دو ساختار اعتقادی را که با توابع جرم m_2, m_1 مشخص شده‌اند در نظر می‌گیریم، اگر C_1 هسته m_1 و C_2 هسته m_2 باشد، آنگاه m_2, m_1 ترکیب‌پذیرند اگر و تنها اگر $C_1 \cap C_2 \neq \varnothing$ برقرار باشد.

دو ساختار اعتقادی را که ترکیب‌ناپذیر باشند، *کاملاً ناسازگار* می‌گویند. معمولاً از عدد $E = \sum_{A\cap B\neq\varnothing} m_1(A) m_2(B) \quad ; \quad 0 \le E \le 1$ به عنوان درجه ناسازگاری یاد می‌شود.

در حالتی‌که E = 1 باشد دو ساختار کاملاً ناسازگار هستند و درحالتی‌که E = 0 باشد دو ساختار کاملاً سازگار با یکدیگر می‌باشند.

حال اگر ساختار اعتقادی با تابع جرم $m(B)=1; \varnothing \neq B \subset \Omega$ را در نظر بگیریم، ترکیب این تابع جرم با هر ساختار اعتقادی دیگر به منزله مشروط‌کردن ساختار اعتقادی به B تعریف می‌شود. می‌توان بقیه روابطی را که در روش دمپستر به دست می‌آوریم، در اینجا هم به دست آورد. کافی است توجه کنیم μ_h در اینجا تبدیل به pls و μ_l تبدیل به bel

می‌شود. پس از آشنایی با قاعده ترکیب دمپستر و شافر، در عمل نیز کارآیی آن را با ذکر یک مثال بررسی می‌کنیم.

مثال: قتلی واقع شده و سه نفر مورد اتهام واقع شده‌اند. مجموعه مرجع Ω به صورت زیر خواهد بود:

$$\Omega = \{Hassan, Ali, Saeed\} = \{H, A, S\}$$

دو شاهد در محضر دادگاه شهادت می‌دهند که توابع جرم جدول 7-2 از نظریات آنها حاصل می‌شود. حال تابع اعتماد را برای هر یک از اعضای کانونی به دست می‌آوریم:

$$bel_1(H \cup A) = \sum_{B \subseteq (H \cup A)} m_1(B) = m_1(H \cup A) + m_1(H) + m_1(A)$$
$$= 0.15 + 0 + 0.05 = 0.2$$

$$bel_2(H \cup A) = \sum_{B \subseteq (A \cup S)} m_2(B) = m_2(A \cup S) + m_2(A) + m_2(S)$$
$$= 0.2 + 0.15 + 0.05 = 0.4$$

با همین روش، دیگر موارد نیز قابل محاسبه می‌باشند.

برای ترکیب شواهد ابتدا باید N را محاسبه نماییم. به عنوان مثال برای $m_{12}(A)$ داریم:

$$\sum_{A \cap B = \varnothing} m_1(A) m_2(B) =$$
$$m_1(A)m_2(H) + m_1(A)m_2(S) + m_1(A)m_2(H \cup S) +$$
$$m_1(H)m_2(A) + m_1(H)m_2(S) + m_1(H)m_2(A \cup S) +$$
$$m_1(S)m_2(A) + m_1(S)m_2(H) + m_1(S)m_2(A \cup H) +$$
$$m_1(A \cup H)m_2(S) + m_1(A \cup S)m_2(H) + m_1(H \cup S)m_2(A)$$
$$= 0.03$$

$$N = 1 - \sum_{A \cap B = \varnothing} m_1(A) m_2(B) = 1 - 0.03 = 0.97$$

حال برای به دست آوردن $m_{12}(A)$ خواهیم داشت:

$$m_{12}(A)=\frac{1}{N}\sum_{B\cap C=A} m_1(B)m_2(C)=$$

$$\left\{\begin{array}{l} m_1(A)m_2(A)+m_1(A)m_2(H\cup S)+m_1(A)m_2(A\cup S) \\ +m_1(A)m_2(H\cup A\cup S)+m_1(H\cup A)m_2(A)+m_1(H\cup A)m_2(H\cup A) \\ +m_1(H\cup A)m_2(A)+m_1(A\cup S)m_2(H\cup A)+m_1(H\cup A\cup S)m_2(A) \end{array}\right\}/0.97=$$

$$\left\{\begin{array}{l} 0.05*0.15+0.05*0.05+0.05*0.2+0.05*0.5+0.15*0.15 \\ +0.15*0.2+0.1*0.15+0.1*0.05+0.6*0.15 \end{array}\right\}/0.97=0.21$$

به صورت مشابه برای پیشامد $A\cup S$ نیز می‌توان نوشت:

$$\begin{aligned} m_{12}(A\cup S) &= \{m_1(A\cup S)m_2(A\cup S)+m_1(A\cup S)m_2(H\cup A\cup S) \\ &+m_1(A\cup S)m_2(A\cup S)\}/0.97 \\ &=0.2 \end{aligned}$$

پس از به دست آوردن m_{12} برای همه اعضای کانونی، می‌توان تابع اعتماد را نیز برای همه آنها محاسبه کرد. به عنوان مثال:

$$\begin{aligned} bel_{12}(H\cup A) &= m_{12}(H\cup A)+m_{12}(H)+m_{12}(A) \\ &=0.12+0.01+0.21=0.34 \end{aligned}$$

نتایج نهایی تمام محاسبات برای توابع جرم و اعتماد در جدول 7-2 به طور کامل بیان شده است:

جدول7-2: محاسبه توابع جرم و اعتماد در مثال داده‌شده

	شاهد 1		شاهد 2		ترکیب شواهد	
اعضا کانون	$m_1(A_i)$	$bel_1(A_i)$	$m_2(A_i)$	$bel_2(A_i)$	$m_{12}(A_i)$	$bel_{12}(A_i)$
H	0	0	0	0	0.01	0.01
A	0.05	0.05	0.15	0.15	0.21	0.21
S	0.05	0.05	0.05	0.05	0.09	0.06
$H\cup A$	0.15	0.2	0.05	0.2	0.12	0.34
$H\cup S$	0.05	0.1	0.05	0.2	0.06	0.16
$A\cup S$	0.1	0.2	0.2	0.4	0.2	0.5

$H \cup A \cup S$	0.6	1	0.5	1	0.3	1

در این فصل، ضمن معرفی تئوری امکان پیشامدهای فازی به تئوری شواهد پرداخته، تفاوت بین مجموعه‌های فازی، اندازه‌های فازی و اندازه‌های احتمال را به تفصیل شرح دادیم. این ایده ممکن است در ذهن خواننده ایجاد شده باشد که می‌توان از تئوری شواهد در سیستمهای خبره استفاده کرد تا بتوان آن را جایگزین انسان خبره نمود.