یکتایی نمایش توزیعهای آمیخته

احمد پارسیان*
عباسعلی خلیلی محمودآبادی†

چکیده

در بسیاری از مسائل کاربردی، تشخیص مدل مناسب برای توزیع صفت جامعهٔ مورد بررسی از اهمیت ویژه‌ای برخوردار است. با توجه به پیشرفتهای اخیر در زمینهٔ کامپیوتر، مدلهای آمیخته به دلیل کاربرد آنها در بسیاری از رشته‌های مختلف علوم بشری از جمله زیست‌شناسی، پزشکی، مهندسی، علوم اجتماعی، ... مورد توجه خاصی قرار گرفته است.

در بررسی مدلهای آمیخته، یکتایی نمایش توزیعهای آمیخته از اهمیت خاصی برخوردار است. در این مقاله سعی شده است که نوشته‌های موجود در این زمینه، به زبانی ساده و به اختصار توضیح داده شود.

مقدمه

فرض کنید یک نمونهٔ تصادفی از قد مردان و زنان یک شهر را در اختیار داریم. ساده‌ترین سؤالی که مطرح می‌شود، این است که:

«چه مدلی برای توزیع متغیر جامعهٔ مورد بررسی مناسب است؟»

می‌دانیم که متوسط قد مردان در یک جامعه معمولاً از متوسط قد زنان بیشتر است. همچنین، با در نظر گرفتن این فرض معقول که توزیع قد مردان و زنان به طور جداگانه هر کدام نرمال هستند، تابع توزیع متغیر مورد نظر، قد جامعه، به صورت زیر خواهد بود:

$$\Phi(x,\mu,\sigma^۲) = p\Phi(x;\mu_۱,\sigma_۱^۲) + (۱-p)\Phi(x;\mu_۲,\sigma_۲^۲)(۱)$$

که در آن ρ نسبت زنان موجود در جامعه و $\Phi(x,\mu_i,\sigma_i^۲)$، $i = ۱, ۲$، به ترتیب توزیع نرمال مربوط به قد زنان و قد مردان در آن جامعه است.

تابع توزیع ارائه شده در (۱)، نمونه‌ای از رده تابعهای توزیع است که از آن تحت عنوان «توزیعهای آمیخته» یاد می‌کنیم.

مسائلی از این قبیل در بسیاری از رشته‌های مختلف علوم، از جمله زیست‌شناسی، پزشکی، مهندسی، همچنین پیشرفتهای کامپیوتری، از دلایل اصلی، مورد توجه قرارگرفتن توزیعهای آمیخته هستند. قبل از ارائهٔ تعریفی دقیق از توزیعهای آمیخته، ارائهٔ چند مثال می‌تواند مفید باشد.

- در زیست‌شناسی، اغلب مشخصه‌های گوناگونی از حیوانات و یا گیاهان، مورد بررسی قرار می‌گیرند که معمولاً تابعی از سن یا جنس حیوان یا گیاه‌اند. از طرفی در اغلب موارد تعیین سن یا جنس حیوان یا گیاه بسیار مشکل است. در نتیجه زیست‌شناس توزیع آمیخته را برای مشخصهٔ مورد بررسی در جامعه درنظر می‌گیرد و عمل آمیختن روی پارامتری که وابسته به متغیر تصادفی غیرقابل مشاهدهٔ سن یا جنس است، انجام می‌گیرد.

- در پزشکی برای مطالعهٔ بیمارانی که دارای فشار خون بالا هستند، یکی از سئوالات موردنظر پژوهشگران علوم پزشکی پاسخگویی به این سؤال است که آیا بیش از یک نوع فشار خون بالا وجود دارد؟ طرح چنین سؤالاتی، منجر به بررسی توزیع آمیخته برای متغیر فشار خون می‌شود.

*) دکتر احمد پارسیان، دانشکدهٔ علوم ریاضی دانشگاه صنعتی اصفهان
†) عباسعلی خلیلی محمودآبادی، دانشکدهٔ علوم ریاضی دانشگاه صنعتی اصفهان

- در مهندسی، یکی از زمینه‌های مهم کاربرد توزیعهای آمیخته، در مطالعهٔ توزیع مشاهدات مربوط به زمانهای خراب شدن قطعات است. اغلب خرابیهای قطعات ممکن است در اثر عوامل متعددی رخ دهد. از طرفی توزیع زمان خرابی مربوط به هر یک از دلایل، می‌تواند توزیع نمایی در نظر گرفته شود. لذا توزیع کلی زمان شکست قطعات، توزیع آمیخته خواهد بود.

با توجه به آنچه اشاره شد، در زیر به تعریف کلی از توزیعهای آمیخته می‌پردازیم.

تعریف ۱. فرض کنید $\mathbb{R}^m_1$ زیرمجموعه‌ای اندازه‌پذیر از فضای اقلیدسی $\mathbb{R}^m$ و مجموعهٔ زیر، خانواده‌ای از توزیعهای مربوط به متغیر تصادفی X باشد:

$$\mathcal{F} = \{F(x,\alpha) : \alpha \in \mathbb{R}^m_1\}.$$

همچنین فرض می‌کنیم $F(x,\alpha)$ یک تابع اندازه‌پذیر روی فضای حاصلضرب x و α باشد. برای هر تابع توزیع mـ‌بعدی و ناتباهیدهٔ G (در این صورت $\mathcal{F}$ حداقل دارای دو عضو است). تابع توزیع H تعریف شده در زیر را «توزیع آمیخته» نسبت به G می‌نامیم.

$$H_G(x) = \int F(x,\alpha)dG(\alpha). \qquad (۲)$$

توزیع H را یک «Gـ آمیخته» از خانوادهٔ $\mathcal{F}$ نیز می‌نامند.

در تعریف فوق می‌توان تابعهای توزیع را با تابعهای چگالی احتمال (یا تابعهای احتمال) متناظر جایگزین کرد. ردهٔ تابعهای توزیع ناتباهیدهٔ G را با $\mathcal{G}$ نشان می‌دهیم و نیز $\mathcal{H}$ و $\mathcal{I}$ را به ترتیب ردهٔ توزیعهای آمیختهٔ H و ردهٔ تابعهای توزیع تباهیده در $\mathbb{R}^m$ در نظر می‌گیریم.

در مطالعهٔ توزیعهای آمیخته، فرض می‌کنیم که فرم تابعی $F(x,\alpha)$ معلوم است و بر اساس مشاهداتی که از تابع توزیع H در اختیار است، هدف برآورد تابع توزیع G است.

در حالتی که توزیع G گسسته و دارای تعداد متناهی پرش است،

$$H(x) = \sum_{i=1}^{k} F(x;\alpha_i)p_i \qquad (۳)$$

را توزیع آمیختهٔ متناهی می‌نامیم و بر اساس مشاهداتی از H، هدف برآورد α_i، p_i و k (اگر مجهول باشد) است.

اگر Gی موجود در تعریف (۲) یکتا نباشد و یا به عبارت دیگر توزیع آمیختهٔ $H(x)$ برای یک خانوادهٔ داده شدهٔ $\mathcal{F}$، دارای نمایش یکتا نباشد، آنگاه در ارتباط با مسائل استنباطی دچار مشکل خواهیم بود. به عنوان مثال، فرض کنید نمونه‌ای از مشاهدات در اختیار است که می‌دانیم از یک توزیع آمیختهٔ متشکل از دو توزیع یکنواخت گرفته شده‌اند. هدف، تجزیه و تحلیل این توزیع برحسب مؤلفه‌های $H(x)$ است. به عبارت دیگر هدف تعیین مؤلفه‌های توزیع آمیخته است. یک روش تحلیل، ممکن است منجر به

$$H(x) = \frac{۱}{۳}u(-۱,۱) + \frac{۲}{۳}u(-۲,۲)$$

و دیگری منجر به

$$H(x) = \frac{۱}{۲}u(-۲,۱) + \frac{۱}{۲}u(-۱,۲)$$

شود. با توجه به اینکه جز نمونهٔ تصادفی، هیچ اطلاع دیگری در مورد مؤلفه‌های توزیع آمیختهٔ مورد بحث در اختیار نیست و روشی وجود ندارد که تعیین کند کدام یک از دو تجزیهٔ فوق درست است، با یک مسألهٔ اساسی روبه‌رو هستیم. بنابراین زمانی که هدف نهایی، دانستن مؤلفه‌های توزیع آمیخته باشد (اگر هدف رده‌بندی مشاهدات آینده برحسب مؤلفه‌های توزیع آمیخته باشد) دچار مشکل خواهیم بود. حتی زمانی که ساختار مؤلفه‌های توزیع آمیخته مهم نباشد، جوابهای چندگانه (یعنی عدم یکتایی G) روشهای برآورد را نیز بسیار مشکل خواهد ساخت.

بنابراین اولین قدم در بررسی و مطالعه توزیعهای آمیخته، بررسی خاصیتی از این توزیعها تحت عنوان «یکتایی نمایش» توزیع آمیخته است. با این مقدمه، در بخش اول خانواده‌هایی از تابعهای توزیع $F(x,\alpha)$ را که توزیع آمیختگی به وجود آمده براساس آنها دارای خاصیت یکتایی نمایش هستند، معرفی می‌کنیم. در بخش دوم شرایط لازم و کافی برای برقراری خاصیت مذکور در مورد آمیخته‌های متناهی را بیان می‌کنیم.

یکتایی نمایش توزیعهای آمیخته

در این بخش تلاش، خواهیم کرد شرایط لازم یا کافی را به اختصار در مورد توزیعهای آمیخته‌ای که می‌توانند «نمایش یکتا» داشته باشند، بیان کنیم. دو حالت خاص پارامتر مقیاس و مکان را بیشتر مورد ارزیابی قرار خواهیم داد. اما ابتدا اشاره به چند تعریف را ضروری می‌دانیم.

تعریف ۲. (خاصیت یکتایی نمایش) گوییم ردهٔ $\mathcal{H}$ در $\mathcal{G}$ (نسبت به $\mathcal{F}$) دارای خاصیت «یکتایی نمایش» است اگر رابطهٔ (۲) یک تناظر یک به یک بین $\mathcal{H} \cup \mathcal{F}$ و $\mathcal{G} \cup \mathcal{I}$ برقرار کند. به عبارت دیگر، اگر

$$H(x) = \int_{\mathbb{R}^m_1} F(x,\alpha)dG(\alpha) = \int_{\mathbb{R}^m_1} F(x,\alpha)dG^*(\alpha),$$

آنگاه برای هر $G^* \in \mathcal{G} \cup \mathcal{I}$،

$$G = G^*.$$

تعریف ۳. گوییم خانوادهٔ توزیعهای $\mathcal{F} = \{F(x;\alpha); \alpha \in D\}$ جمعی است اگر

$$F(x;\alpha_1) * F(x;\alpha_2) = F(x;\alpha_1 + \alpha_2);$$

به عبارت دیگر اگر X_1 و X_2 دو متغیر تصادفی مستقل به ترتیب با توزیعهای $F(x;\alpha_1)$ و $F(x;\alpha_2)$ باشند، آنگاه $X_1 + X_2$ دارای تابع توزیع $F(x;\alpha_1 + \alpha_2)$ است.

تعریف ۴. مجموعهٔ D را یک نیم گروه آبلی ـ جمعی گوییم، هرگاه نسبت به عمل جمع دارای خاصیت شرکتپذیری و عضو خنثی باشد.

مجموعه‌های $D(I^+)$، $D(Q^+)$ و $D(\mathbb{R}^+)$ را به ترتیب به عنوان نیم‌گروههای آبلی ـ جمعی اعداد صحیح مثبت، گویای مثبت، و حقیقی مثبت در نظر می‌گیریم.

با توجه به تعاریف ارائه شده، اگر $\mathcal{F} = \{F(x;\alpha) : \alpha \in D\}$ یک خانواده از توزیعهای جمعی، همچنین $F(x;\alpha)$ یک تابع اندازه‌پذیر روی فضای حاصلضرب x و α، و D یک زیرمجموعهٔ اندازه‌پذیر از فضای اقلیدسی $\mathbb{R}^m$ باشد، آنگاه

$$H(x) = \int_D F(x;\alpha) dG(\alpha)$$

یک تابع توزیع آمیخته از خانوادهٔ جمعی $\mathcal{F}$ را بیان می‌کند. به سادگی می‌توان نشان داد که اگر $m = 1$، آنگاه تابع مشخصهٔ $F(x;\alpha)$ به صورت زیر است:

$$\Phi(t;\alpha) = [\Phi(t)]^\alpha, \alpha \geq 0$$

که در آن $\Phi(t) = \Phi(t;1)$.

قضیهٔ زیر شرایط کافی برای «یکتایی نمایش» توزیعهای آمیختهٔ ساخته شده براساس خانوادهٔ توزیعهای جمعی را بیان می‌کند.

قضیهٔ ۱: اگر $m = 1$ و D یکی از مجموعه‌های $D(I^+)$، $D(Q^+)$، یا $D(\mathbb{R}^+)$ باشد، آنگاه خانوادهٔ توزیعهای $\mathcal{H}$ از آمیخته‌های خانوادهٔ توزیعهای جمعی $\mathcal{F}$، دارای خاصیت «یکتایی نمایش» است. (نگاه کنید به [۶]).

خانوادهٔ توزیعهای جمعی زیر، مثالهایی از به کارگیری قضیهٔ ۱ هستند، به عبارت دیگر بر اساس قضیهٔ ۱ توزیعهای آمیختهٔ ساخته شده بر اساس هر یک از خانوادهٔ توزیعهای زیر دارای خاصیت «یکتایی نمایش» هستند.

$$\{N(0,\sigma^2) : \sigma^2 > 0\},$$
$$\{\Gamma(r,\lambda) : r > 0, \text{معلوم } \lambda\},$$
$$\{\chi^2_{(k)} : k > 0\},$$
$$\{P(\lambda) : \lambda > 0\},$$
$$\{B(n,p) : n \in \mathbb{N}, \text{معلوم } p\},$$
$$\{NB(r,p) : r \in \mathbb{N}, \text{معلوم } p\}.$$

تذکر یک نکته در اینجا ضروری است.

- نتیجهٔ قضیهٔ ۱ را برای $m > 1$ نمی‌توان تعمیم داد. در مجموعهٔ مثالهایی که ارائه خواهیم کرد، نشان می‌دهیم که در توزیع نرمال وقتی که هر دو پارامتر آن تحت یک توزیع دو بعدی تغییر می‌کنند، توزیع آمیخته ساخته شده براساس آن دارای خاصیت «یکتایی نمایش» نیست.

قضیهٔ ۲: فرض کنید $m = 1$ و $\mathcal{F}$ خانواده‌ای از توزیعهای جمعی باشد. اگر

الف) $\phi(t;\alpha)$ تابعی حقیقی از t باشد و برای یک t_0 (متناهی یا نامتناهی) $\lim_{t \to t_0} \phi(t;\alpha) = 0$، یا

ب) $\mathcal{F}$ دارای گشتاورهای دوم متناهی و گشتاور اول مخالف صفر باشد، آنگاه هیچ توزیع آمیخته‌ای از این خانواده متعلق به $\mathcal{F}$ نخواهد بود. (نگاه کنید به [۵]).

در ادامهٔ این بخش، به بررسی خاصیت «یکتایی نمایش» در خانوادهٔ توزیعهای مقیاس، یعنی $\mathcal{F}_1 = \{F(x;\alpha) = F(x\alpha)\}$ و خانواده توزیعهای مکان، یعنی $\mathcal{F}_\alpha = \{F(x;\alpha) = F(x-\alpha)\}$ می‌پردازیم. این امر به این دلیل است که حالتهایی از این دو خانواده وجود دارند ($\Gamma(r,\lambda)$، r معلوم و λ مجهول و $N(\theta,1)$) که نتایج قضایای فوق برای خاصیت «یکتایی نمایش» در مورد آنها صدق نمی‌کند.

خانواده توزیعهای مقیاس

فرض کنید $m = 1$، آنگاه توزیع آمیخته بر پایهٔ خانوادهٔ $\mathcal{F}_1$ به صورت زیر خواهد بود

$$H(x) = \int_0^\infty F(x\alpha) dG(\alpha) \qquad (4)$$

به طوری که $F(0^+) = 0$. توزیع F متعلق به $\mathcal{F}_1$ را اصطلاحاً «توزیع مولد» می‌نامند. در زیر نشان می‌دهیم اگر $F(e^x)$ دربازه‌ای از اعداد حقیقی دارای تبدیل فوریه مخالف صفر باشد، آنگاه توزیعهای آمیخته روی خانوادهٔ $\mathcal{F}_1$ دارای خاصیت «یکتایی نمایش» اند.

برای این کار، با انتخاب $x = e^y$ و $\alpha = e^{-\beta}$، $\bar{F}_0(\omega) = F_0(e^\omega)$ و $\bar{G}(\beta) = 1 - G(e^{-\beta})$، از رابطهٔ (۴) نتیجه می‌شود که

$$\bar{H}(y) = H(e^y)$$

$$
\begin{aligned}
&= \int_{+\infty}^{-\infty} F(e^{y-\beta})\,dG(e^{-\beta}) \\
&= -\int_{-\infty}^{+\infty} \bar{F}(y-\beta)\,d\{1-\bar{G}(\beta)\} \\
&= \int_{-\infty}^{+\infty} \bar{F}(y-\beta)\,d\bar{G}(\beta)
\end{aligned}
$$

بنابراین،

$$\bar{H}(y) = \bar{F} * \bar{G}.$$

حال فرض کنید که

$$
\begin{aligned}
H(x) &= \int_0^{\infty} F(x\alpha)\,dG_1(\alpha) \\
&= \int_0^{\infty} F(x\alpha)\,dG_2(\alpha).
\end{aligned}
$$

با توجه به آنچه گفته شد، معلوم می‌شود

$$\bar{H}(y) = \bar{F} * \bar{G}_1 = \bar{F} * \bar{G}_2. \qquad (۵)$$

حال اگر از طرفین رابطهٔ (۵) تبدیل فوریه بگیریم، داریم

$$
\begin{aligned}
\psi_{\bar{H}}(t) &= \int_{-\infty}^{+\infty} e^{ity}\,d\bar{H}(y) \\
&= \int_{-\infty}^{\infty} e^{ity}\,d(\bar{F} * \bar{G}_1)(y) \\
&= \int_{-\infty}^{\infty} e^{ity}\,d(\bar{F} * \bar{G}_2)(y).
\end{aligned}
$$

بنابراین،

$$\psi_{\bar{F}}(t).\psi_{\bar{G}_1}(t) = \psi_{\bar{F}}(t).\psi_{\bar{G}_2}(t). \qquad (۶)$$

مشاهده می‌کنیم که اگر $\bar{F}(x) = F(e^x)$ در بازه‌ای از اعداد حقیقی دارای تبدیل فوریهٔ مخالف صفر باشد، آنگاه از رابطهٔ (۶) نتیجه می‌شود که $\bar{G}_1 = \bar{G}_2$ ؛ یعنی $G_1 = G_2$.

خانواده توزیعهای مکان

فرض کنید $m = 1$، آنگاه توزیع آمیخته بر پایهٔ خانوادهٔ $\mathcal{F}_2$ به صورت زیر خواهد بود

$$
\begin{aligned}
H(x) &= \int F(x-\alpha)\,dG(\alpha) \\
&= (F * G)(x).
\end{aligned}
$$

حال اگر

$$\int F(x-\alpha)\,dG_1(\alpha) = \int F(x-\alpha)\,dG_2(\alpha),$$

آنگاه

$$F * G_1 = F * G_2;$$

یعنی

$$\psi_F(t)\psi_{G_1}(t) = \psi_F(t)\psi_{G_2}(t).$$

بنابراین در این حالت نیز اگر توزیع F روی بازه‌ای از اعداد حقیقی دارای تبدیل فوریهٔ مخالف صفر باشد، آنگاه $G_1 = G_2$.

در پایان این بخش، به تفصیل، مثالهای متنوعی را ارائه می‌کنیم.

مثال ۱: خانواده توزیعهای گاما با پارامترهای γ و λ را در نظر بگیرید. در این صورت

$$F(x;\lambda,\gamma) = \gamma^{\lambda}[\Gamma(\lambda)]^{-1} \int_0^x u^{\lambda-1} e^{-\gamma u}\,du$$

و تابع مشخصهٔ آن به صورت $\phi(t;\lambda,\gamma) = [1 - (it/\gamma)]^{-\lambda}$ است.

- اگر γ ثابت باشد، آنگاه این خانواده یک خانواده از توزیعهای جمعی است و بنابر قضیهٔ ۱، ردهٔ توزیعهای آمیخته بر پایهٔ این خانواده دارای خاصیت «یکتایی نمایش» است.
- اگر λ ثابت باشد، آنگاه این خانواده، یک خانواده از توزیعهای مقیاس γ است، که دارای تابع «توزیع مولد» $F(x) = F(x;\lambda,1)$ است، به عبارت دیگر $F_0(x\gamma) = F(x;\lambda,\gamma)$.

به سادگی معلوم می‌شود که تابع مشخصهٔ $\bar{F}(x)$ به صورت زیر است

$$
\begin{aligned}
\psi_{\bar{F}}(t) &= \int_{-\infty}^{\infty} e^{itx}\,d\bar{F}(x) \\
&= \frac{\Gamma(\lambda+it)}{\Gamma(\lambda)}.
\end{aligned}
$$

حال چون $\psi_{\bar{F}}(t)$ مخالف صفر است، بنابراین ردهٔ توزیعهای آمیخته بر پایهٔ توزیع گاما با λی ثابت، دارای خاصیت «یکتایی نمایش» است.

- در سال ۱۹۴۹، رابینز نشان داد، زمانی که هر دو پارامتر λ و γ تحت توزیع پیوستهٔ $G(\lambda,\gamma)$ تغییر کنند، آنگاه توزیعهای آمیخته دارای خاصیت «یکتایی نمایش» نیستند.

مثال ۲: خانواده توزیعهای یکنواخت با پارامترهای θ و σ را در نظر بگیرید. در این صورت

$$
F(x;\theta,\sigma) = \begin{cases} 0 & x < \theta - \sigma \\ \frac{x-\theta+\sigma}{2\sigma} & \theta - \sigma \le x < \theta + \sigma, \theta \in \mathbb{R}, \sigma > 0 \\ 1 & \theta + \sigma \le x \end{cases}
$$

توزیع آمیخته بر اساس خانوادهٔ $\{F(x;\theta,\sigma) : \theta \in \mathbb{R}, \sigma \in \mathbb{R}^+\}$ به صورت زیر است:

$$\begin{aligned} H(x) &= \int F(x;\theta,\sigma)dG(\theta,\sigma) \\ &= \int F(x;\theta,\sigma)dG_1(\theta)dG_\theta(\sigma) \end{aligned}$$

به طوری که $G_\theta(\sigma)$ توزیع شرطی σ به شرط θ و $G_1(\theta)$ توزیع حاشیه‌ای θ است.

- اگر σ ثابت باشد ($\sigma = 1$)، آنگاه خانوادهٔ توزیعها، متعلق به خانوادهٔ توزیعهای مکان است و «توزیع مولد» و تابع چگالی احتمال مولد آن به صورت زیر است

$$F(x) = \begin{cases} 0 & x < -1 \\ \frac{x+1}{2} & -1 \le x < 1 \\ 1 & 1 \le x \end{cases}$$

$$f(x) = \frac{1}{2} \qquad -1 < x < 1.$$

به سادگی معلوم می‌شود که تبدیل فوریهٔ توزیع F برابر است با

$$\psi_F(t) = \int_{-1}^{1} e^{itx} f(x)dx = \frac{\sin t}{t}.$$

با توجه به اینکه $\psi_F(t)$ دارای تعداد شمارش‌پذیری ریشهٔ حقیقی است، می‌توان بازه‌هایی از اعداد حقیقی پیدا کرد که در آنها $\psi_F(t)$ مخالف صفر است. بنابراین با توجه به آنچه گفته شد، نتیجه می‌شود که توزیع آمیخته $H(x) = \int F(x;\theta,1)dG_1(\theta)$ دارای خاصیت «یکتایی نمایش» است.

- اگر پارامتر θ ثابت باشد ($\theta = 0$)، آنگاه خانوادهٔ توزیعها، متعلق به خانوادهٔ توزیعهای مقیاس است و توزیع مولد و تابع چگالی احتمال مولد آن به ترتیب به صورت زیر است

$$F(x) = \begin{cases} 0 & x < -\sigma \\ \frac{x}{2\sigma} & -\sigma \le x < \sigma \\ 1 & \sigma \le x \end{cases}$$

و

$$f(x) = \frac{1}{2\sigma} \qquad -\sigma < x < \sigma.$$

بنابراین با توجه به اینکه

$$h(x) = \int_{|x|}^{+\infty} (2\sigma)^{-1} dG(\sigma),$$

نتیجه می‌شود که

$$dh(x) = -(2x)^{-1} dG(|x|)$$

یا

$$dh(\sigma) = -(2\sigma)^{-1} dG(\sigma). \qquad (7)$$

حال اگر دو توزیع G_1 و G_2 یک چگالی آمیخته h را تولید بکنند، آنگاه با توجه به رابطهٔ (۷)،

$$dG_1(\sigma) = dG_2(\sigma),$$

یعنی

$$G_1(\sigma) = G_2(\sigma) + k.$$

اما با توجه به اینکه $G_1(0^+) = G_2(0^+) = 0$، بنابراین $k = 0$، یعنی توزیعهای آمیخته بر اساس خانوادهٔ توزیعهای $\{F(x;0,\sigma)\}$ دارای خاصیت «یکتایی نمایش» است.

- زمانی که هر دو پارامتر θ و σ در توزیع $F(x;\theta,\sigma)$ تغییر کنند، خانواده توزیعهای آمیخته بر پایهٔ این خانواده دارای خاصیت «یکتایی نمایش» نیست: برای مثال

$$\begin{aligned} \frac{1}{3}F(x;0,1,) + \frac{2}{3}F(x;0,2) &= \frac{1}{2}F(x;-\frac{1}{2},\frac{3}{2}) \\ &+ \frac{1}{2}F(x;\frac{1}{2},\frac{3}{2}) \end{aligned}$$

اما $G_1 \neq G_2$، یا

$$F(x;\frac{1}{2},\frac{1}{2}) = \frac{1}{2}F(x;\frac{1}{4},\frac{1}{4}) + \frac{1}{2}F(x;\frac{1}{2},\frac{3}{2}),$$

اما مجدداً $G_1 \neq G_2$.

با توجه به مثالهای ارائه شده، ممکن است تصور شود که ردهٔ توزیعهای آمیخته براساس خانوادهٔ توزیعهای یک پارامتری همواره دارای خاصیت «یکتایی نمایش» است. مثال زیر این موضوع را روشن می‌کند.

مثال ۳: خانوادهٔ ترزیعهای دوجمله‌ای با پارامترهای n و θ را درنظر بگیرید که در آن n ثابت است. تابع احتمال توزیعهای آمیخته براساس این خانواده به صورت زیر است

$$\begin{aligned} h(x) &= \int_0^1 \binom{n}{x} \theta^x (1-\theta)^{n-x} dG(\theta) \\ &= \int_0^1 \binom{n}{x} \theta^x \{\sum_{k=0}^{n-x} \binom{n-x}{k} (-\theta)^k\} dG(\theta) \\ &= \sum_{k=0}^{n-x} \binom{n}{x} \binom{n-x}{k} (-1)^k \int_0^1 \theta^{x+k} dG(\theta) \\ &= \sum_{k=0}^{n-k} a_{x,k} m_{x+k}(G), \quad x = 0, 1, \ldots, n \end{aligned}$$

که در آن

$$a_{x,k}=\binom{n}{x}\binom{n-x}{k}(-1)^k$$

و

$$m_{x+k}(G)=\int_0^1\theta^{x+k}dG(\theta).$$

با توجه به مقادیر ممکن x، معلوم می‌شود که $h(x)$ ترکیبی خطی از n گشتاور اول توزیع G است. با توجه به اینکه می‌توان دو توزیع متفاوت G_1 و G_2 با n گشتاور اول برابر پیدا کرد، نتیجه می‌شود که $h(x)$ دارای خاصیت «یکتایی نمایش» نیست.

مثال ۴: فرض کنید $\mathcal{F}=\{P(x;\theta):\theta\in(0,1)\}$ از خانوادهٔ تابعهای احتمال زیر باشد

$$p(x;\theta)=P_\theta(X=x)=\begin{cases}\theta & x=1\\ (1-\theta)^2\theta^{x-2} & x=2,3,\dots\end{cases}$$

نشان دهید خانوادهٔ توزیعهای آمیخته براساس خانوادهٔ $\mathcal{F}$ دارای خاصیت «یکتایی نمایش» است.

حل:

$$\begin{aligned}p(x;G)&=p(x;G)\\&=\int_0^1 p(x;\theta)dG(\theta)\\&=\begin{cases}m_1(G), & x=1\\ m_{x-2}(G)-2m_{x-1}(G)+m_x(G), & x=2,3,\dots\end{cases}\end{aligned}$$

اگر برای هر $x=1,2,\dots$، $p(x;G_1)=p(x;G_2)$، آنگاه

$$\begin{aligned}x=1&:m_1(G_1)=m_1(G_2)\\ x=2&:m_0(G_1)-2m_1(G_1)+m_2(G_1)\\&=m_0(G_2)-2m_1(G_2)+m_2(G_2).\end{aligned}$$

که از آن نتیجه می‌شود که

$$m_1(G_1)=m_2(G_2).$$

$$\begin{aligned}x=3&:m_1(G_1)-2m_1(G_1)+m_2(G_1)\\&=m_3(G_2)-2m_4(G_2)+m_0(G_2)\\ \vdots\quad&\quad\vdots\end{aligned}$$

که از آن نتیجه می‌شود که

$$m_3(G_1)=m_3(G_2).$$

بنابراین با ادامهٔ روند فوق به نتیجهٔ زیر خواهیم رسید:

$$m_i(G_1)=m_i(G_2),\ i=1,2,\dots$$

حال با استفاده از این نتیجه که بین یک تابع توزیع تعریف شده روی بازهٔ $[0,1]$ و دنبالهٔ گشتاورهای آن تناظر یک به یک وجود دارد (مرجع [۲])، بنابراین

$$G_1=G_2.$$

- تذکر یک نکته در اینجا ضروری است.
- اگر برای عضوی از خانوادهٔ مورد بررسی $\mathcal{F}$، مانند $F(x;\beta)$، رابطهٔ زیر برقرار باشد

$$H(x)=F(x;\beta)=\int F(x,\alpha)dG(\alpha)\qquad(8)$$

(توجه داشته باشید، که بنا به تعریف، G یک توزیع ناتباهیده است)، آنگاه رابطهٔ (۸) دلیل بر عدم برقراری خاصیت «یکتایی نمایش» برای این خانواده است. زیرا به ازای یک توزیع تباهیدهٔ G_1 در β که مخالف G است، داریم

$$H(x)=\int F(x;\alpha)dG_1(\alpha)=\int F(x;\alpha)dG(\alpha).$$

از طرف دیگر عدم برقراری خاصیت «یکتایی نمایش» برای یک خانوادهٔ $\mathcal{F}$، لزوماً دلیل بر این نیست که بتوان برای عضوی از این خانواده، رابطهٔ (۸) را نوشت. به عنوان مثال، مجدداً خانواده توزیعهای دوجمله‌ای با $n\geq 2$ (ثابت) و θ را در نظر گیرید. می‌دانیم که این خانواده دارای خاصیت «مکتایی نمایش» نیست. نشان می‌دهیم که برای هیچ عضوی از این خانواده نمی‌توان رابطهٔ (۸) را برقرارکرد. این کار را با برهان خلف به اثبات می‌رسانیم. فرض کنید رابطهٔ (۸) برای عضوی از خانواده برقرار باشد، آنگاه

$$\binom{n}{x}\theta_0^x(1-\theta_0)^{n-x}=\int_0^1\binom{n}{x}\theta^x(1-\theta)^{n-x}dG(\theta).$$

بنابراین

$$\begin{aligned}\int_0^1&\binom{n}{x}\theta^x(1-\theta)^{n-x}dG_1(\theta)\\&=\int_0^1\binom{n}{x}\theta^x(1-\theta)^{n-x}dG(\theta),\end{aligned}$$

که در آن

$$G_1(\theta)=\begin{cases}0 & \theta<\theta_0\\ 1 & \theta\geq\theta_0.\end{cases}$$

با توجه به مثال ۳، n گشتاور اول G و G_1 باید با هم برابر باشند. از طرفی چون G_1 یک توزیع تباهیده است، گشتاور اول و دوم آن برابر $\theta_\circ$ و $\theta_\circ^2$ است. بنابراین باید

$$\theta_\circ = \int_\circ^1 \theta dG(\theta), \qquad \theta_\circ^2 = \int_\circ^1 \theta^2 dG(\theta),$$

یعنی $V_G(\theta) = 0$، که نشان دهندهٔ تباهیده بودن توزیع G در $\theta_\circ$ است و این خلاف فرض است. بنابراین هیچ عضوی از خانوادهٔ توزیع دوجمله‌ای را نمی‌توان به صورت رابطهٔ (۸) نوشت، علی‌رغم آنکه توزیع .سیختهٔ ساخته شده بر اساس آن خاصیت «یکتایی نمایش» را ندارد.

در سال ۱۹۴۹، رابینز نشان داد که خانوادهٔ توزیعهای آمیخته بر اساس توزیع گاما با پارامترهای λ و γ (هر دو متغیر) دارای خاصیت «یکتایی نمایش» نیستند. در زیر نشان می‌دهیم که خانوادهٔ توزیعهای آمیخته بر اساس توزیع $N(\theta, \sigma^2)$ نیز دارای خاصیت «یکتایی نمایش» نیستند.

در این حالت، توزیع آمیخته تعریف شده در (۲) به صورت زیر است

$$\begin{aligned} H(x) &= \int_{\mathbb{R}^2} \Phi(x; \theta, \sigma^2) dG(\theta, \sigma^2) \\ &= \int_{\mathbb{R}^2} \Phi\left(\frac{x-\theta}{\sigma}\right) dG(\theta, \sigma^2) \\ &= \int_{\mathbb{R}^2} \Phi\left(\frac{x-\theta}{\sigma}\right) dG_{\sigma^2}(\theta) dG_1(\sigma^2) \qquad (۹) \end{aligned}$$

به طوری که $G_1(\sigma^2)$ توزیع مربوط به σ^2 و $G_{\sigma^2}(\theta)$ توزیع شرطی θ به شرط σ^2 است. بنابراین

$$\begin{aligned} H(x) &= \int_\circ^\infty \left\{ \int_{-\infty}^\infty \Phi\left(\frac{x-\theta}{\sigma}\right) dG_{\sigma^2}(\theta) \right\} dG_1(\sigma^2) \\ &= \int_\circ^\infty \left\{ \Phi(x; 0, \sigma^2) * G_{\sigma^2}(x) \right\} dG_1(\sigma^2). \qquad (۱۰) \end{aligned}$$

با تبدیل فوریه گرفتن از طرفین رابطهٔ (۱۰) داریم

$$\varphi(t) = \int_\circ^\infty e^{-\sigma^2 t^2/2} \varphi_{\sigma^2}(t) dG_1(\sigma^2).$$

حال اگر $H(x)$ متقارن باشد، $\varphi(t)$ حقیقی خواهد بود. در این حالت نشان می‌دهیم که $H(x)$ دارای خاصیت «یکتایی نمایش» نیست.

فرض کنید برای مقادیر ثابت $\theta_\circ$ و $\sigma_\circ^2$،

$$H(x) = \Phi(x; \theta_\circ, \sigma_\circ^2) = \Phi\left(\frac{x-\theta_\circ}{\sigma_\circ}\right).$$

حال اگر $H(x)$ یک G^*ـ آمیخته از خانواده $\{\Phi(x; \theta, \sigma^2)\}$ باشد، آنگاه

$$\Phi\left(\frac{x-\theta_\circ}{\sigma_\circ}\right) = \int_{\mathbb{R}^2} \Phi\left(\frac{x-\theta}{\sigma}\right) dG^*(\theta, \sigma^2),$$

که با تبدیل x به $\sigma_\circ x + \theta_\circ$ داریم

$$\begin{aligned} \Phi(x) &= \int_{\mathbb{R}^2} \Phi\left(\frac{x - \frac{\theta-\theta_\circ}{\sigma_\circ}}{\sigma/\sigma_\circ}\right) dG^*(\theta, \sigma^2) \\ &= \int_{\mathbb{R}^2} \Phi\left(\frac{x-\theta}{\sigma}\right) dG(\theta, \sigma^2). \qquad (۱۱) \end{aligned}$$

بنابراین بدون اینکه از کلیت مسأله کم شود، می‌توان فرض کرد $\sigma_\circ = 1$ و $\theta_\circ = 0$. با تبدیل لاپلاس گرفتن از طرفین رابطهٔ (۱۱) داریم

$$e^{\frac{s^2}{2}} = \int_{\mathbb{R}^2} e^{\frac{\sigma^2 s^2}{2} - \theta s} dG(\theta, \sigma^2). \qquad (۱۲)$$

حال اگر طرفین رابطهٔ (۱۲) را در $e^{\frac{\beta s^2}{2}}$ ضرب کنیم، داریم

$$e^{(\beta+1)s^2/2} = \int_{\mathbb{R}^2} e^{(\sigma^2+\beta)\frac{s^2}{2} + \theta s} dG(\theta, \sigma^2). \qquad (۱۳)$$

با تبدیل s به $s(1+\beta)^{-\frac{1}{2}}$ و σ^2 به $\sigma^2(1+\beta) - \beta$ و θ به $\theta(1+\beta)^{\frac{1}{2}}$ در رابطهٔ (۱۳) داریم

$$\begin{aligned} e^{s^2/2} &= \int_{\mathbb{R}^2} e^{\sigma^2 s^2/2 + \theta s} dG(\theta\sqrt{1+\beta}, \sigma^2(1+\beta) - \beta) \\ &= \int_{\mathbb{R}^2} e^{\sigma^2 s^2/2 + \theta s} dG_\beta(\theta, \sigma^2). \qquad (۱۴) \end{aligned}$$

بنابراین با توجه به روابط (۱۱)، (۱۲) و (۱۴) مشاهده می‌کنیم که $\Phi(x)$ نه تنها یک Gـ آمیخته از خانوادهٔ $\{\Phi(x; \theta, \sigma^2)\}$ است، بلکه یک G_βـ آمیخته از خانوادهٔ $\{\phi(x; \theta, \sigma^2)\}$ نیز هست. لذا دارای خاصیت «یکتایی نمایش» نیست.

یک حالت خاص نیز با انتخاب $G_{\sigma^2}(x) = \Phi(x; 0, 1-\sigma^2)$ در رابطهٔ (۱۰) به سادگی قابل مشاهده است. در این صورت مقدار داخل آکولاد برابر $\Phi(x)$ است که بستگی به σ^2 ندارد و برای هر توزیع $G_1(\sigma^2)$ روی $(0, 1)$، داریم $H(x) = \Phi(x)$، که نشان می‌دهد خاصیت «یکتایی نمایش» برقرار نیست.

آمیخته‌های متناهی

در توزیع تعریف شده در رابطهٔ (۲)، هرگاه G یک توزیع گسسته با جرمهای احتمال مثبت در تعدادی متناهی نقطه از فضای $\mathbb{R}^m$ باشد، آنگاه H را یک توزیع آمیخته متناهی می‌نامیم. در این بخش، به دلیل اهمیت این نوع توزیعها، به بررسی خاصیت «یکتایی نمایش» این گونه توزیعهای آمیخته می‌پردازیم. فرض کنید $\mathcal{G}_k$، $k = 1, 2, \ldots$ ردهٔ توزیعهای گسسته‌ای باشد

که دارای جرم احتمال مثبت در دقیقاً k نقطه هستند و $\mathcal{G}=\bigcup_{k=1}^{\infty}\mathcal{G}_k$. خانواده توزیعهای آمیختهٔ تولید شده به وسیلهٔ این ردهها را به ترتیب با $\mathcal{H}_k$، $k=1,2,\ldots$ و $\mathcal{H}$ نمایش میدهیم. بنابراین

$$\mathcal{H}=\{H(x):H(x)=\sum_{i=1}^{k}p_iF(x;\alpha_i),\,F(x;\alpha)\in\mathcal{F},$$

$$p_i>0,\,i=1,\ldots,k,\sum_{i=1}^{k}p_i=1,\,k=1,2,\ldots\}.$$

خاصیت «یکتایی نمایش» در آمیختههای متناهی را میتوان به صورت زیر بیان کرد: اگر داشته باشیم،

$$\sum_{i=1}^{k_1}p_iF(x;\alpha_i)=\sum_{i=1}^{k_2}p_i'F(x;\alpha_i'),$$

آنگاه نتیجه بگیریم که $k_1=k_2$ و برای هر $1\leq i\leq k_1$، jیی در بازهٔ ۱ تا $k_1=k_2$ وجود دارد به طوری که $p_i=p_j'$ و $F(x;\alpha_i)=F(x;\alpha_j')$.

برای سهولت در نمادهای به‌کار رفته از این به بعد $F(x;\alpha_i)$ را با $F_i(x)$ نمایش میدهیم. قضیهٔ زیر شرط لازم و کافی برای «یکتایی نمایش» را در خانوادهٔ $\mathcal{H}$ بیان میکند. (نگاه کنید به [۳]).

قضیهٔ ۳: شرط لازم و کافی برای اینکه ردهٔ $\mathcal{H}$ از آمیختههای متناهی خانوادهٔ $\mathcal{F}$ دارای خاصیت «یکتایی نمایش» باشد، آن است که عناصر $\mathcal{F}$ روی مجموعهٔ اعداد حقیقی، مستقل خطی باشد.

برای نتیجه‌گیری از قضیهٔ ۳، آگاهی از تعریف زیر ضروری است. فرض کنید V و W دو فضای برداری باشند.

تعریف ۵: تبدیل $T:V\longrightarrow W$ را خطی گوییم اگر برای هر $v_1,v_2\in V$ و مقادیر ثابت a,b،

$$T(av_1+bv_2)=aT(v_1)+bT(v_2).$$

در صورتی که تبدیل T یک به یک و پوشا باشد، آن را **یکریختی** میگوییم.

نتیجهٔ ۱: شرط لازم و کافی برای اینکه $\mathcal{H}$ دارای خاصیت «یکتایی نمایش» باشد آن است که تصویر $\mathcal{F}$ تحت هر یکریختی روی فضای $<\mathcal{F}>$ (فضای تولید شده توسط خانوادهٔ $\mathcal{F}$) مستقل خطی باشد. (نگاه کنید به [۳]).

حال فرض کنید $\phi(t)$ تبدیلی خطی و یک به یک روی خانوادهٔ توزیعهای $\mathcal{F}$ باشد، که دارای دامنهٔ تعریف S_ϕ است. عملگر > بین دو توزیع F_1,F_2 از $\mathcal{F}$ را به صورت زیر تعریف میکنیم.

«$F_1<F_2$ اگر و فقط اگر (۱) $S_{\phi_1}\subseteq S_{\phi_2}$ و (۲) برای t_1ای متعلق به $\bar{S}_{\phi_1}$ و مستقل از ϕ_2 وجود داشته باشد که $\lim_{t\to t_1}\frac{\phi_2(t)}{\phi_1(t)}=0$». در حقیقت عملگر > یک نوع ترتیب بین اعضای $\mathcal{F}$ برقرار میکند.

قضیهٔ ۴: در صورت برقراری عملگر فوق بین عناصر خانوادهٔ توزیعهای $\mathcal{F}$، خانوادهٔ توزیعهای $\mathcal{H}$ دارای خاصیت «یکتایی نمایش» است. (نگاه کنید به [۷]).

مثالهای زیر کاربرد قضیهٔ ۴ را در بررسی «یکتایی نمایش» نشان میدهند.

مثال ۵: فرض کنید

$$\mathcal{F}=\{N(\theta,\sigma^2):\theta\in(-\infty,\infty),\sigma\in(0,\infty)\}.$$

خانواده توزیعهای آمیخته متناهی $\mathcal{H}$ از خانوادهٔ $\mathcal{F}$ دارای خاصیت یکتایی نمایش است.

حل: فرض کنید $\Phi(x;\theta,\sigma^2)\equiv\Phi$ نمایانگر تابع توزیع نرمال با میانگین θ و واریانس σ^2 باشد. تبدیل لاپلاس دو طرفهٔ چگالی Φ برابر است با

$$\begin{aligned}\phi(t;\theta,\sigma^2)&=\int_{-\infty}^{\infty}e^{-tx}d\Phi(x;\theta,\sigma^2)\\&=e^{\frac{\sigma^2t^2}{2}-\theta t},\qquad S_\phi=(-\infty.+\infty).\end{aligned}$$

بنابراین

$$\frac{\phi_2(t;\theta_2,\sigma_2^2)}{\phi_1(t;\theta_1,\sigma_1^2)}=e^{\frac{t^2}{2}(\sigma_2^2-\sigma_1^2)-t(\theta_2-\theta_1)}$$

و

$$\lim_{t\to+\infty}\frac{\phi_2(t;\theta_2,\sigma_2^2)}{\phi_1(t;\theta_1,\sigma_1^2)}=\begin{cases}0 & \sigma_2<\sigma_1\\ \text{یا} & \\ 0 & \sigma_1=\sigma_2\ \&\ \theta_1<\theta_2\end{cases}$$

با توجه به اینکه امکان مرتب کردن عناصر $\mathcal{F}$ با توجه به عملگر تعریف شده وجود دارد، خانوادهٔ $\mathcal{H}$ دارای خاصیت «یکتایی نمایش» است.

مثال ۶: فرض کنید $\mathcal{F}=\{\Gamma(\alpha,\theta):\alpha>0,\theta>0\}$. خانوادهٔ توزیعهای آمیختهٔ متناهی $\mathcal{H}$ از خانوادهٔ $\mathcal{F}$ دارای خاصیت «یکتایی نمایش» است.

حل: فرض کنید $F(x;\alpha,\theta)$ نمایانگر تابع توزیع $\Gamma(\alpha,\theta)$ باشد. تبدیل لاپلاس چگالی F برابر است با

$$\begin{aligned}\phi(t;\alpha,\theta)&=\int_0^{\infty}e^{-tx}dF(x;\alpha,\theta)\\&=(1+\frac{t}{\theta})^{-\alpha};\qquad S_\phi=(-\infty,+\infty).\end{aligned}$$

بنابراین

$$\begin{aligned}\frac{\phi_2(t)}{\phi_1(t)} &= \frac{\phi(t;\alpha_2,\theta_2)}{\phi(t;\alpha_1,\theta_1)}\\ &= \frac{(1+\frac{t}{\theta_1})^{\alpha_1}}{(1+\frac{t}{\theta_2})^{\alpha_2}}\end{aligned}$$

و

$$\lim_{t\to-\theta_1}\frac{\phi(t;\alpha_2,\theta_2)}{\phi(t;\alpha_1,\theta_1)} = \begin{cases} 0, & \theta_1 < \theta_2 \\ \text{یا} & \\ 0, & \theta_1 = \theta \& \alpha_2 < \alpha_1. \end{cases}$$

حال با توجه به اینکه امکان مرتب کردن عناصر $\mathcal{F}$ با توجه به عملگر تعریف شده وجود دارد، خانوادهٔ $\mathcal{H}$ دارای خاصیت «یکتایی نمایش» است.

تذکر یک نکته در اینجا ضروری است.

با توجه به آنچه در ارتباط با توزیع نرمال و گاما در بخش ۱ و مثالهای ۵ و ۶ عنوان شد، معلوم می‌شود که هرگاه در رابطهٔ (۲)، G یک توزیع پیوسته دوبعدی باشد، آنگاه $H(x)$ دارای خاصیت «یکتایی نمایش» نیست ولی اگر G یک توزیع گسستهٔ دوبعدی با تعداد نقاط متناهی جرم احتمال مثبت باشد، آنگاه $H(x)$ دارای خاصیت «یکتایی نمایش» است.

مثال ۷: فرض کنید

$$\mathcal{F} = \{B(n_i,p_i) : n_i \neq n_j, i \neq j, 0 < p_i < 1, i = 1,2,\dots\}.$$

نشان دهید خانوادهٔ توزیعهای آمیخته متناهی $\mathcal{H}$ از خانوادهٔ $\mathcal{F}$ دارای خاصیت «یکتایی نمایش» است.

حل: تبدیل لاپلاس توزیع $B(n_i,p_i)$ برابر است با

$$\phi(t;n,p) = (q+pe^{-t})^n,\ q = 1-p, S_\phi = (-\infty,+\infty)$$

بنابراین

$$\frac{\phi_2(t)}{\phi_1(t)} = \frac{\phi(t;n_2,p_2)}{\phi(t;n_1,p_1)} = \frac{(q_2+p_2e^{-t})^{n_2}}{(q_1+p_1e^{-t})^{n_1}}.$$

و

$$\lim_{t\to-\infty}\frac{\phi(t;n_2,p_2)}{\phi(t;n_1,p_1)} = 0,\ n_1 > n_2.$$

حال با توجه به اینکه امکان مرتب کردن عناصر $\mathcal{F}$ با توجه به عملگر تعریف شده وجود دارد، خانوادهٔ $\mathcal{H}$ دارای خاصیت «یکتایی نمایش» است.

مثال ۸: فرض کنید $\mathcal{F} = \{NB(r,p) : 0 < r, 0 < p < 1\}$. نشان دهید خانوادهٔ توزیعهای آمیختهٔ متناهی $\mathcal{H}$ از خانوادهٔ $\mathcal{F}$ دارای خاصیت «یکتایی نمایش» است.

حل: تابع احتمال هر عضو از خانوادهٔ $\mathcal{F}$ به صورت زیر است:

$$f(x;r,p) = \binom{r+x-1}{x}p^rq^x,\ x = 0,1,2,\dots,\ r > 0,\quad 0 < p < 1$$

تبدیل تابع مولد احتمال توزیع $NB(r,p)$ برابر است با

$$\begin{aligned}\phi(t;r,p) &= E(t^X)\\ &= \sum_{x=0}^{\infty} t^x\binom{r+x-1}{x}p^rq^x\\ &= \left(\frac{p}{1-qt}\right)^r,\ |qt| < 1.\end{aligned}$$

بنابراین

$$S_\phi = \{t : |t| < (1-p)^{-1}\}$$

$$\frac{\phi_2(t)}{\phi_1(t)} = \frac{\phi(t;r_2,p_2)}{\phi(t;r_1,p_1)} = \frac{(1-q_1t)^{r_1}p_2^{r_2}}{(1-q_2t)^{r_2}p_1^{r_1}}$$

و

$$\lim_{t\to-(1-p_2)^{-1}}\frac{\phi_2(t)}{\phi_1(t)} = \begin{cases} 0, & p_1 < p_2 \\ \text{یا} & \\ 0, & p_1 = p_2 \& r_1 < r_2. \end{cases}$$

حال با توجه به اینکه امکان مرتب کردن عناصر $\mathcal{F}$ با توجه به عملگر تعریف شده وجود دارد، خانوادهٔ توزیعهای آمیختهٔ $\mathcal{H}$ دارای خاصیت «یکتایی نمایش» است.

دو مثال بعدی کاربردِ قضیهٔ ۳ و نتیجهٔ ۱ را در بررسی «یکتایی نمایش» نشان می‌دهند.

مثال ۹: فرض کنید $\mathcal{F} = \{C(\alpha,\beta) : \alpha \in \mathbb{R}, \beta > 0\}$، که در آن $C(\alpha,\beta)$ نمایانگر توزیع کوشی با پارامتر مکان α و پارامتر مقیاس β است. نشان دهید خانوادهٔ توزیعهای آمیخته متناهی $\mathcal{H}$ از خانوادهٔ $\mathcal{F}$ دارای خاصیت «یکتایی نمایش» است.

حل: تابع چگالی احتمال توزیع $C(\alpha,\beta)$ برابر است با

$$f(x;\alpha,\beta) = \frac{\beta}{\pi[\beta^2+(x-\alpha)^2]}, x \in \mathbb{R}, \beta > 0, \alpha \in \mathbb{R}.$$

با انتخاب M_1 به عنوان عملگر تابع مشخصه داریم

$$\begin{aligned} M_1(f(x;\alpha,\beta)) &= \exp\{i\alpha t - \beta|t|\} \\ &= \phi(t;\alpha,\beta),\ t \in \mathbb{R}. \end{aligned}$$

عملگر خطی M_2 روی تابعهای ϕ را به صورت زیر اختیار می‌کنیم

$$\begin{aligned} M_2(\phi(t;\alpha,\beta)) &= \begin{cases} \phi(t;\alpha,\beta), & t \geq 0 \\ 0, & t < 0 \end{cases} \\ &= \psi(t;\alpha,\beta). \end{aligned}$$

بالاخره عملگر خطی M_3 روی تابعهای ψ را تبدیل لاپلاس اختیار می‌کنیم، یعنی

$$\begin{aligned} M_3(\psi(t;\alpha,\beta)) &= M_3M_2M_1(f(x;\alpha,\beta)) \\ &= \int_0^{+\infty} e^{-st}e^{-\beta t + i\alpha t}dt \\ &= \int_0^{+\infty} e^{-t(s+\beta-i\alpha)}dt \\ &= \int_0^{+\infty} e^{-t(s+b)}dt \end{aligned}$$

که در آن $b = \beta - i\alpha$. بنابراین

$$M_3(\psi(t;\alpha,\beta)) = \frac{1}{s+b},\ s \in D_\beta$$

به طوری که

$$D_\beta = \{x + iy : x > -\beta\}.$$

توجه کنید که تبدیل $M_3M_2M_1$ یک یکریختی روی $\mathcal{F}$ است و تصویر رابطهٔ خطی

$$0 = \sum_{j=1}^{N} d_j f(x;\alpha_j,\beta_j)$$

تحت این تبدیل به صورت زیر خواهد بود

$$0 = \sum_{j=1}^{N} d_j(s+b_j)^{-1}$$

که در آن $s \in D = \bigcap_{j=1}^{N} D_{\beta_j}$ و $b_j = \beta_j - i\alpha_j$. فرض کنید که برای $j > 1$، $\beta_1 \leq \beta_j$. با ضرب کردن طرفین رابطهٔ فوق در $s + b_1$ خواهیم داشت،

$$0 = d_1 + \sum_{j=2}^{N} d_j(s+b_j)^{-1}(s+b_1)$$

یا

$$d_1 = -\sum_{j=2}^{N} d_j(s+b_j)^{-1}(s+b_1).$$

بنابراین

$$|d_1| \leq |s+b_1| \sum_{j=2}^{N} |d_j(s+b_j)^{-1}|$$

و

$$|d_1| \leq \lim_{\substack{s \to -b_1 \\ s \in D}} |s+b_1| \sum_{j=2}^{N} |d_j(s+b_j)^{-1}| = 0.$$

در نتیجه $d_1 = 0$. حال با تکرار فرایند فوق معلوم می‌شود که

$$d_j = 0, \quad j = 1, \ldots, N;$$

یعنی تصویر $\mathcal{F}$ تحت یکریختی $M_3M_2M_1$، مستقل خطی است. لذا بنابر نتیجهٔ ۱، عناصر $\mathcal{F}$ نیز مستقل خطی هستند و خانوادهٔ توزیعهای آمیخته تولید شده از $\mathcal{F}$ دارای خاصیت «یکتایی نمایش» است.

مثال ۱۰: فرض کنید

$$\mathcal{F} = \{F(x;\alpha) : F(x;\alpha) = F(x-\alpha), \alpha \in \mathbb{R}\}.$$

نشان دهید خانوادهٔ توزیعهای آمیختهٔ متناهی $\mathcal{H}$ از خانوادهٔ $\mathcal{F}$ دارای خاصیت «یکتایی نمایش» است.

حل: با انتخاب M_1 به عنوان عملگر تابع مشخصه داریم

$$\begin{aligned} M_1(F(x;\alpha)) = \phi_\alpha(t) &= \int e^{itx}dF(x-\alpha) \\ &= \int e^{it(y+\alpha)}dF(y) \\ &= e^{it\alpha}\int e^{ity}dF(y) \\ &= e^{it\alpha}\phi_0(t) \\ &= M_1(F(x))e^{it\alpha}. \end{aligned}$$

تصویر $\mathcal{F}$ تحت تبدیل M_1 را با نماد $M_1(\mathcal{F})$ نشان می‌دهیم. تصویر رابطهٔ خطی

$$0 = \sum_{j=1}^{N} C_j F(x;\alpha_j) \qquad (15)$$

تحت این تبدیل به صورت زیر خواهد بود

$$
\begin{aligned}
۰ &= \sum_{j=۱}^{N} C_j \phi_{\alpha_j}(t) \\
&= \left(\sum_{j=۱}^{N} C_j e^{it\alpha_j}\right) \phi_{۰}(t).
\end{aligned}
$$

با توجه به اینکه $\phi_{۰}(t)$ یک تابع مشخصه است، داریم

$$\lim_{t\to ۰} \phi_{۰}(t) = \phi_{۰}(۰) = ۱.$$

بنابراین dی مثبتی وجود دارد که برای هر $t \in (-d, d)$،

$$۰ = \sum_{j=۱}^{N} C_j e^{it\alpha_j}.$$

از طرف دیگر با توجه به قضیه‌ای (مرجع [۱]، ص ۲۸۴) نتیجه می‌شود که

$$۰ = \sum_{j=۱}^{N} C_j e^{it\alpha_j} \qquad \forall t \in \mathbb{R}.$$

اگر عملگر $M_۲$ را تبدیل لاپلاس در نظر گیریم، آنگاه تصویر رابطهٔ خطی (۱۵) تحت تبدیل $M_۲M_۱$ به صورت زیر خواهد بود،

$$
\begin{aligned}
۰ &= \sum_{j=۱}^{N} C_j \int_{۰}^{\infty} e^{-t(s-i\alpha_j)} dt \\
&= \sum_{j=۱}^{N} C_j (s-i\alpha_j)^{-۱}, \qquad Re(s) > ۰.
\end{aligned}
$$

حال با ضرب طرفین رابطهٔ فوق در $(s - i\alpha_k)$، داریم

$$C_k + \sum_{j\neq k} C_j (s-i\alpha_j)^{-۱}(s-i\alpha_k) = ۰, Re(s) > ۰.$$

بنابراین، برای $Re(s) > ۰$

$$C_k = -\sum_{j\neq k} C_j (s-i\alpha_j)^{-۱}(s-i\alpha_k)$$

و

$$|C_k| \le |s-i\alpha_k| \sum_{j\neq k} |C_j||s-i\alpha_j|^{-۱}.$$

حال اگر $s \to i\alpha_k$، آنگاه، $|C_k| \le ۰$، یعنی $C_k = ۰$. این فرایند برای هر k، $k = ۱, \ldots, N$ درست است، بنابراین برای هر k، $k = ۱, \ldots, N$، $C_k = ۰$.

از طرف دیگر چون تبدیل $M_۲M_۱$ یک یکریختی روی فضای $< \mathcal{F} >$ است، بنابراین با استفاده از نتیجهٔ ۱، آمیختهٔ متناهی به دست آمده براساس خانوادهٔ توزیعهای $\mathcal{F}$، دارای خاصیت «یکتایی نمایش» است.

مراجع

[1] Churchill, Ruel Vance. (1974). Complex variables and applications.

[2] Feller, W. (1971). An introduction to probability theory and its applications, II. Wiley, New York.

[3] Sidney J. Yakowitz and John D. Spragins. (1968) On the identifiability of finite mixtures. Ann. Math. Stat. Vol. 39, No. 1, 209-214.

[4] Teicher, H. (1953). On the convolution of distributions. Ann. Math. Stat. 775-778.

[5] Teicher, H. (1960). On the mixture of distributions. Ann. Math. Stat. 31, 55-73.

[6] Teicher, H. (1961). Identitiability of mixtures. Ann. Math. Stat. 32, 244-248.

[7] Teicher, H. (1963). Identifiability of finite mixtures. Ann. Math. Stat. 34, 1265-1269.