توزیع‌های آمیخته متناهی با مولفه های نرمال چوله

محمد بهرامی [۱]

چکیده:

یک توزیع آمیخته ترکیبی از دو یا چند توزیع آماری است. وقتی نمونه گیری از یک جامعه غیر همگن متشکل از دو یا چند زیرجامعه هر یک با توزیعی متفاوت انجام می‌گیرد، با توزیعی آمیخته سر و کار پیدا می‌کنیم. برای مثال، توزیع زمان خرابی آمیخته‌ای از مؤلفه‌های سالم و معیوب، توزیع وزن حیوانات با گروه‌های سنی متفاوت یا مدت زمانی که بیماران قلبی پس از یک عمل جراحی در گروههای سنی متفاوت زندگی خواهند کرد. درچنین حالاتی نمودار منحنی مربوط به تابع چگالی اندازه خصوصیت مورد نظر در جامعه آماری می‌تواند دارای چند مد یا نما با شد. همچنین در بسیاری از موارد با خصوصیت‌هایی مواجه می شویم که مجموعه مقادیر آنها می تواند متعلق به مجموعه اعداد حقیقی باشد و در واقع چون این‌گونه خصوصیت‌ها معمولا به طور طبیعی پدیدار می‌گردند، به نظر می‌رسد که توزیع نرمال برای آنها مناسب باشد. اما با مطالعه بیشتر در می‌یابیم که وضعیت خصوصیت مورد مطالعه برای افراد جامعه کاملا متقارن نبوده بلکه به راست یا به چپ دارای چولگی می‌باشد. یعنی تابع چگالی مورد نظر دارای چولگی به راست یا به چپ می‌باشد. بنابراین متغیر هایی با مجموعه مقادیر اعداد حقیقی دارای عدم تقارن مورد نظر خواهد بود. این موضوع می‌تواند منجر به تعریف توزیعی به نام توزیع نرمال چوله گردد. در این مقاله حالتی را بررسی می‌کنیم که مؤلفه‌های یک توزیع آمیخته از توزیعهای نرمال چوله تشکیل شده‌اند. بنابراین ابتدا توزیع آمیخته و سپس توزیع نرمال چوله را بررسی می‌کنیم و در پایان توزیع آمیخته با مؤلفه‌های نرمال را تعریف و برآورد پارامترهای موجود در این مدل را با استفاده از الگوریتم $EM-$ به دست می آوریم.

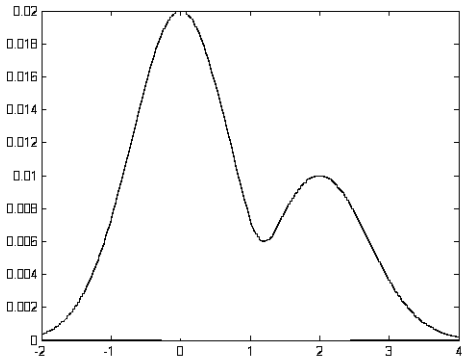
واژه‌های کلیدی: الگوریتم $EM-$ ، تابع درستنمایی، توزیع آمیخته، توزیع نرمال چوله، چگالی آمیخته، متغیرهای پنهان.

۱ مقدمه

در بسیاری موارد با رسم منحنی فراوانی نسبی داده‌ها و مشاهدهٔ دو یا چند نقطهٔ ماکسیمم در آنها، پی می‌بریم که جامعه آماری از لحاظ خصوصیت مورد بررسی یک جامعه ناهمگن [۲] می‌باشد. نمودار فراوانی نسبی زیر با توجه به دونمایی [۳] بودن آن حاکی از نا همگنی جامعه آماری مورد نظر است. این نمودار در واقع می‌تواند معرف

[۱] استادیار گروه آمار دانشگاه اصفهان
[۲] Non-Homogeneous Population
[۳] Bimodal

یک تابع چگالی آمیخته باشد.



شکل ۱. نمودارتابع چگالی آمیخته با دو نما.

۲ معرفی توزیع آمیخته

فرض کنید جامعه آماری P شامل k $(k \geq ۲)$ زیر جامعه $P_۱, P_۲, \ldots, P_k$ باشد. هدف بررسی خصوصیتی در این جامعه است. همچنین فرض کنید متغیر تصادفی X نشان دهنده مقدار خصوصیت مذکور باشد. اگر $F_۱, F_۲, \ldots, F_k$ تابع توزیع متغیر تصادفی X به ترتیب در زیر جامعه‌های $P_۱, P_۲, \ldots, P_k$ باشند، آنگاه تابع توزیع متغیر تصادفی X با تکیه گاه S در جامعه P به صورت زیر خواهد بود:

$$F(x) = \pi_۱ F_۱(x) + \pi_۲ F_۲(x) + \ldots + \pi_k F_k(x), x \in S \quad (۱)$$

که در آن π_i احتمال آن است که X دارای توزیع F_i باشد. $۰ < \pi_i < ۱$ برای $i = ۱, ۲, \ldots, k$ و $\sum_{i=۱}^{k} \pi_i = ۱$ و مقادیر π_i ها را نسبت‌های آمیختگی می‌نامند. مدل (۱) یک توزیع آمیخته متناهی با k مؤلفه بر حسب تابع توزیع است. چنانچه $f_۱, f_۲ \ldots, f_k$ توابع چگالی متغیر تصادفی در زیر جامعه‌های $P_۱, P_۲, \ldots, P_k$ و f تابع چگالی احتمال X در جامعه P باشد، آنگاه مدل آمیخته را می‌توان به صورت زیر بر حسب تابع چگالی نوشت:

$$\begin{aligned} f(x) &= \pi_۱ f_۱(x) + \ldots + \pi_k f_k(x) \\ &= \sum_{i=۱}^{k} \pi_i f_i(x), x \in S \end{aligned} \quad (۲)$$

در حالت کلی لزومی ندارد که مؤلفه‌های $f_۱, f_۲, \ldots, f_k$ متعلق به یک خانواده از چگالی‌ها باشند. توابع چگالی f_i برای $i = ۱, ۲, \ldots, k$ می‌توانند شامل بردارپارامتر $\underline{\theta}_i = (\theta_{i۱}, \theta_{i۲}, \ldots, \theta_{ih})$ باشند. بنابراین یک چگالی آمیخته متناهی پارامتری با k مؤلفه به شکل کلی زیر است:

$$f(x|\psi) = \sum_{i=۱}^{k} \pi_i f(x|\underline{\theta}_i), \underline{\theta}_i \in \Theta, x \in S \quad (۳)$$

که در آن $\psi = (\pi_۱, \pi_۲, \ldots, \pi_k, \underline{\theta}'_۱, \underline{\theta}'_۲, \ldots, \underline{\theta}'_k)' \in \Psi$ و ψ را بردارپارامتر مدل آمیخته می‌نامیم که در فضای پارامتر Ψ تغییر می کند.

۳ نمونه‌گیری از یک جامعه ناهمگن

فرض کنید $X_۱, X_۲, \ldots, X_n$ نمونه‌ای تصادفی به اندازه n از جامعه ناهمگن P باشد. واضح است هر یک از نمونه‌های X_j با احتمال π_j متعلق به زیر جامعه P_i است. بنابراین برای هر مشاهده X_j داریم:

$$f(x_j|\psi) = \sum_{i=۱}^{k} \pi_i f(x_j|\underline{\theta}_i), \quad \underline{\theta}_i \in \Theta, x \in S \quad (۴)$$

برای مشخص نمودن اینکه نمونه X_j متعلق به زیرجامعه P_i می‌باشد، از یک متغیر تصادفی کمکی استفاده می‌شود که به آن متغیر پنهان [۴] می‌گوییم.

[۴] Latent Variable

اگر این متغیرها را با Z_j نشان دهیم، آنگاه برای نمونه: $X_1, X_2, ..., X_n$ بردار $Z_j = (Z_{1j}, Z_{2j}, ..., Z_{kj})$ $(j = 1, ..., n)$ به صورت زیر تعریف می‌شود:

$$Z_{ij} = \begin{cases} 1 & \text{اگر } X_j \text{ متعلق به زیر جامعه } P_i \text{ باشد} \\ 0 & \text{در غیر این‌صورت} \end{cases} \quad (۵)$$

با به‌کارگیری متغیرهای پنهان Z_j، داده‌های کامل به صورت زیر تعریف می‌شود:

$$\{Y_j; j = 1, 2, ..., n\} = \{(X_j, Z_j); j = 1, 2, ..., n\} \quad (۶)$$

واضح است که $\sum_{j=1}^{n} Z_{ij} = 1$ برای $i = 1, 2, ..., k$ و داریم:

$$\sum_{i=1}^{k} z_{ij} f(x_j|\theta_i) = \prod_{i=1}^{k} f(x_j|\theta_i)^{z_{ij}} \quad (۷)$$

از سوی دیگر $(Z_j|\underline{\theta}, \underline{\pi})$ برای $j = 1, 2, ..., k$ از هم مستقل بوده و دارای توزیع چند جمله ای به صورت زیر است.

$$(Z_j|\underline{\theta}, \underline{\pi}) \sim M(1, \pi_1, ..., \pi_k)$$

بنا براین داریم:

$$f(z_j|\underline{\theta}, \underline{\pi}) = \frac{1!}{z_{1j}! z_{2j}! ... z_{kj}!} \prod_{i=1}^{k} \pi_i^{z_{ij}} \sim \prod_{i=1}^{k} \pi_i^{z_{ij}} \quad (۸)$$

همچنین طبق (۷) داریم:

$$f(x_j|z_j, \underline{\theta}, \underline{\pi}) = \sum_{i=1}^{k} z_{ij} f(x_j|\theta_i) = \prod_{i=1}^{k} f(x_j|\theta_i)^{z_{ij}} \quad (۹)$$

بنابراین می‌توان نوشت:

$$\begin{aligned} f(x_j, z_j|\underline{\theta}, \underline{\pi}) &= f(x_j|z_j, \underline{\theta}, \underline{\pi}) f(z_j|\underline{\theta}, \underline{\pi}) \\ &= \prod_{i=1}^{k} [\pi_i f(x_j|\theta_i)]^{z_{ij}} \quad (۱۰) \end{aligned}$$

در نتیجه برای نمونه تصادفی $X_1, X_2, ..., X_n$ تابع درستنمایی در داده‌های کامل به صورت زیر خواهد بود:

$$L(\psi|y_1, y_2, ..., y_n) = L(\psi|\underline{x}, \underline{z}) = \prod_{j=1}^{n} f(x_j, z_j|\psi) \quad (۱۱)$$

و با به‌کارگیری (۱۰) داریم:

$$L(\psi|y_1, y_2, ..., y_n) = \prod_{j=1}^{n} \prod_{i=1}^{k} [\pi_i \ f(x_j|\underline{\theta}_i)]^{Z_{ij}} \quad (۱۲)$$

لگاریتم $L(\psi)$ به صورت زیر نوشته می‌شود:

$$\begin{aligned} &\log L(\psi|y_1, y_2, ..., y_n) = l(\psi|y_1, y_2, ..., y_n) \\ &= \sum_{j=1}^{n} \sum_{i=1}^{k} z_{ij} \log \pi_i + \sum_{j=1}^{n} \sum_{i=1}^{k} z_{ij} \ \log f(x_j|\underline{\theta}_i) \quad (۱۳) \end{aligned}$$

یکی از مسائل اصلی در توزیع‌های آمیخته، مسئله برآورد پارامترها می‌باشد. معمولا برآوردهای گشتاوری، برآوردهای دقیقی نبوده و در بسیاری از مواقع این برآوردها بسیار نادرست و غیر واقعی هستند. اگر از (۱۳) نسبت به پارامترهای موجود مشتق گرفته و مساوی صفر قرار دهیم، می‌توانیم برآوردهای ماکسیمم درستنمایی این پارامترها را به دست آوریم. از آنجا که حل معادلات بوجود آمده توسط مشتق گیری در توزیعهای آمیخته کاری دشوار است بنابراین باید برای به دست آوردن برآورد پارامترها از روشهای عددی استفاده کنیم. از معروفترین روشهای عددی موجود در به دست آوردن برآوردهای ماکسیمم درستنمایی پارامترها می‌توان الگوریتم EM—[۵] را نام برد که توسط دمپستر [۲] معرفی گردید.

[۵] Expectation and Maximization Algorithm

۴ توزیع نرمال چوله

فرض کنید X یک متغیر تصادفی با مجموعه مقادیر $\mathbb{R}$ باشد، اگر تابع چگالی این متغیر تصادفی به صورت:

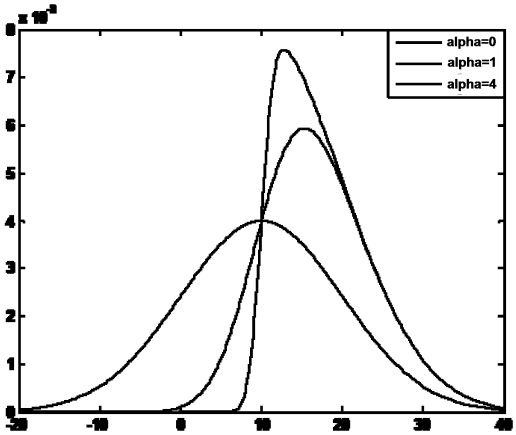
$$f(x|\mu,\sigma^2,\alpha)=\frac{2}{\sigma}\phi(\frac{x-\mu}{\sigma})\Phi(\alpha\frac{x-\mu}{\sigma}) \quad (۱۴)$$

باشد، در این صورت می‌گوییم X دارای توزیع نرمال چوله[۶] با پارامترهای μ,σ^2 و α بوده و آن را با نماد: $X\sim SN(\mu,\sigma^2,\alpha)$ نشان می‌دهیم. این توزیع اولین بار توسط آزالینی [۱] معرفی گردید. در (۱۴)، ϕ و Φ به ترتیب تابع چگالی و توزیع نرمال استاندارد می‌باشند. α را پارامتر شکل (Shape Parameter) می‌نامیم. به عبارت دیگر تابع چگالی نرمال چوله را می‌توان به صورت زیر نشان داد.

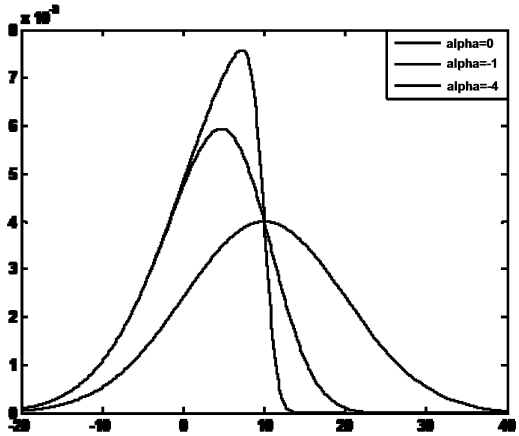
$$\begin{aligned} f(x|\mu,\sigma^2,\alpha) &= \frac{2}{\sigma}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\{-\frac{1}{2\sigma^2}(x-\mu)^2\} \\ &\times \int_{-\infty}^{\alpha(\frac{x-\mu}{\sigma})}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz \end{aligned}$$

باید دقت داشته باشیم هرگاه: $\alpha=0$ باشد در این‌صورت: $f(x|\mu,\sigma^2,\alpha=0)=\phi(\frac{x-\mu}{\sigma})$ که نشان دهنده توزیع نرمال با پارامترهای μ و σ^2 است. همچنین اگر α مثبت باشد چولگی به سمت راست و در صورتی که α منفی باشد چولگی نمودار تابع چگالی به سمت چپ خواهد بود. چنانچه $\alpha\to\infty$ آن‌گاه تابع چگالی به سمت تابع چگالی نیم— نرمال (Half-Normal) میل خواهد کرد. نمودارهای زیر وضعیت چولگی تابع چگالی نرمال چوله را به ازای مقادیر مختلف نشان می‌دهد.

[۶] Skew Normal Distribution



شکل ۲. نمودار تابع چگالی نرمال چوله به ازای مقادیر مثبت $(\mu=10,\sigma=10)$.



شکل ۳. نمودار تابع چگالی نرمال چوله به ازای مقادیر منفی $(\mu=10,\sigma=10)$.

چنانچه در (۱۴) $\mu=0$ و $\sigma^2=1$ باشد، در این صورت تابع چگالی نرمال چوله استاندارد را خواهیم داشت. و در این حالت تنها پارامتر موجود، پارامتر چولگی بوده و تابع چگالی آن را به صورت: $f(z|\alpha)=2\phi(z)\Phi(\alpha z)$ نشان می‌دهیم. پیسی [۴] در مورد توزیع نرمال معرفی شده توسط آزالینی [۱] مسائلی را حل نمود. رینالدو [۳] نیز

یک حالت کلی تر از توزیع نرمال چوله را که به آن توزیع نرمال چوله تعمیم یافته[۷] گفته می‌شود معرفی کرد که تابع چگالی آن به صورت زیر تعریف می‌شود.

$$f(x|\alpha_1,\alpha_2)=2\phi(x)\Phi(\frac{\alpha_1 x}{\sqrt{1+\alpha_2 x^2}});x\in R,\alpha_1\in R,\alpha_2\geq 0$$

لازم به ذکر است که تابع چگالی فوق مربوط به توزیع نرمال استاندارد چوله تعمیم یافته با پارامترهای α_1 و α_2 بوده و آن را با نماد $X\sim SGN(0,1,\alpha_1,\alpha_2)$ نشان می‌دهیم. با تبدیل: $Y=\mu+\sigma X$ شکل کلی تابع چگالی نرمال چوله تعمیم یافته به صورت زیر به دست می آید.

$$f(y|\mu,\sigma,\alpha_1,\alpha_2)=\frac{2}{\sigma}\phi(\frac{y-\mu}{\sigma})\Phi\{\frac{\alpha_1(y-\mu)}{\sqrt{\sigma^2+\alpha_2(y-\mu)^2}}\}$$

برای به دست آوردن تابع چگالی نرمال چوله می‌توانیم ازقضیه (۱) استفاده کنیم. اما ابتدا تعریف زیر را خواهیم داشت.

تعریف ۴–۱ فرض کنید Y یک متغیر تصادفی باشد. می‌گوییم این متغیر تصادفی دارای توزیع نرمال بریده شده[۸] در فاصله (a_1,a_2) می‌باشد و آن را با نماد $Y\sim TN(\mu,\sigma^2)I\{a_1<y<a_2\}$ نشان می‌دهیم، هرگاه تابع چگالی آن به صورت زیر باشد.

$$f(y|\mu,\sigma^2)=\left\{\Phi(\frac{a_2-\mu}{\sigma})-\Phi(\frac{a_1-\mu}{\sigma})\right\}^{-1}\times\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\};a_1<y<a_2$$

قضیه ۴–۱ فرض کنید متغیر تصادفی U دارای توزیع نرمال با میانگین صفر و واریانس σ^2 و متغیر تصادفی T دارای توزیع بریده شده نرمال به صورت: $T\sim TN(0,\sigma^2)I\{t>0\}$ باشد. اگر U و T مستقل باشند، آنگاه متغیرتصادفی: $X=\mu+\frac{\alpha}{\sqrt{1+\alpha^2}}T+\frac{1}{\sqrt{1+\alpha^2}}U$ دارای توزیع نرمال چوله با پارامترهای μ, σ^2 و α خواهد بود. **اثبات:** با استفاده از تبدیل زیر می توان قضیه را اثبات نمود.:

$$\begin{cases}X=\mu+\frac{\alpha}{\sqrt{1+\alpha^2}}T+\frac{1}{\sqrt{1+\alpha^2}}U\\ W=T\end{cases}$$

نتیجه ۴–۱ با توجه به روابط فوق داریم:

$$T|X=x\sim TN(\delta(\alpha)(x-\mu),\sigma^2(1-\delta^2(\alpha)))I\{t>0\}\qquad(15)$$

فرع ۴–۱ فرض کنید متغیر تصادفی Y دارای توزیع بریده شده با پارامترهای μ و σ^2 درفاصله (a_1,a_2) باشد. در اینصورت می‌توان نشان داد:

$$E(Y)=\mu-\sigma\frac{\phi(\frac{a_2-\mu}{\sigma})-\phi(\frac{a_1-\mu}{\sigma})}{\Phi(\frac{a_2-\mu}{\sigma})-\Phi(\frac{a_1-\mu}{\sigma})}$$

و

$$E(Y^2)=\mu^2+\sigma^2-2\mu\sigma\frac{\phi(\frac{a_2-\mu}{\sigma})-\phi(\frac{a_1-\mu}{\sigma})}{\Phi(\frac{a_2-\mu}{\sigma})-\Phi(\frac{a_1-\mu}{\sigma})}-\sigma^2\frac{(\frac{a_2-\mu}{\sigma})\phi(\frac{a_2-\mu}{\sigma})-(\frac{a_1-\mu}{\sigma})\phi(\frac{a_1-\mu}{\sigma})}{\Phi(\frac{a_2-\mu}{\sigma})-\Phi(\frac{a_1-\mu}{\sigma})}$$

بنابراین با استفاده از (۱۵) و فرع (۴–۱) داریم:

$$E(T|X=x)=\delta(\alpha)(x-\mu)+\frac{\phi(\frac{\delta(\alpha)(x-\mu)}{\sigma\sqrt{1-\delta^2(\alpha)}})}{\Phi(\frac{\delta(\alpha)(x-\mu)}{\sigma\sqrt{1-\delta^2(\alpha)}})}\times\sigma\sqrt{1-\delta^2(\alpha)}\qquad(16)$$

[۷] Skew- Generalized Normal Distribution
[۸] Truncated Normal Distribution

$$E(T^{۲}|X=x)=\delta^{۲}(\alpha)(x-\mu)^{۲}+\sigma^{۲}(۱-\delta^{۲}(\alpha))$$
$$+\frac{\phi(\frac{\delta(\alpha)(x-\mu)}{\sigma\sqrt{۱-\delta^{۲}(\alpha)}})}{\Phi(\frac{\delta(\alpha)(x-\mu)}{\sigma\sqrt{۱-\delta^{۲}(\alpha)}})}\delta(\alpha)(x-\mu)\sigma^{۲}(۱-\delta^{۲}(\alpha)) \quad (۱۷)$$

در اینجا نیز باید دقت داشته باشیم که برآورد ماکسیمم درستنمایی برای پارامترها در توزیع نرمال چوله به صورت دقیق میسر نبوده و باید از روشهای عددی مانند الگوریتم $EM-$ استفاده کنیم.

۵ تابع چگالی آمیخته با مولفه‌های نرمال چوله

فرض کنید یک متغیر تصادفی با تابع چگالی آمیخته زیر باشد:

$$f(x|\psi)=\sum_{i=۱}^{k}\pi_i g(x|\mu_i,\sigma_i^{۲},\alpha_i);x\in R \quad (۱۸)$$

به طوری که $g(x|\mu_i,\sigma_i^{۲},\alpha_i)$ نشان دهنده تابع چگالی نرمال چوله با پارامترهای μ_i، $\sigma_i^{۲}$ و پارامتر شکل α_i است. بردار پارامتر ψ به صورت: $\psi=(\underline{\theta}_۱,\underline{\theta}_۲,...,\underline{\theta}_k,\pi_۱,\pi_۲,...,\pi_k)$ بوده که در آن: $\underline{\theta}_i=(\mu_i,\sigma_i^{۲},\alpha_i)$ می‌باشد. فرض کنید نمونه تصادفی $X_۱,X_۲,...,X_n$ از x انتخاب و $Z_j=(Z_{۱j},Z_{۲j},...,Z_{kj})$ بردار متغیرهای پنهان باشد. همچنین فرض کنید $T_۱,T_۲,...,T_n$ نیز یک نمونه تصادفی از توزیع نرمال بریده شده در فاصله $(۰,+\infty)$ باشد. بنابراین طبق (۱۵) داریم:

$$T_j|x_j,z_{ij}=۱\sim TN(\mu_{T_{ij}},\sigma_{T_i}^{۲})I\{t_j>۰\}$$
$$;j=۱,۲,...,n \quad (۱۹)$$
$$T_j|z_{ij}=۱\sim TN(۰,\sigma_i^{۲})I\{t_j>۰\} \quad (۲۰)$$
$$Z_j\sim M(۱,\pi_۱,\pi_۲,...,\pi_k);j=۱,۲,...,n \quad (۲۱)$$

که $\mu_{T_{ij}}=\delta(\alpha_i)(x_j-\mu_i)$ و $\sigma_{T_i}=\sigma_i\sqrt{۱-\delta^{۲}(\alpha_i)}$، $i=۱,۲,...,k$. اگر داده‌های کامل را $(\underline{X},\underline{T},\underline{Z})$ در نظر بگیریم، در این صورت مدل سلسله مراتبی[۹] برای داده‌های کامل به صورت زیر است:

$$f(x_j,t_j,z_j|\psi)=f(x_j,t_j|z_j,\psi)f(z_j|\underline{\pi}) \quad (۲۲)$$

که در (۲۲) داریم:

$$\begin{aligned}f(x_j,t_j|z_{ij},\psi) &= \sum_{i=۱}^{k}z_{ij}f(x_j,t_j|\psi)\\ &= \prod_{i=۱}^{k}[f(x_j,t_j|\psi)]^{z_{ij}}\end{aligned}$$

و یا

$$f(x_j,t_j|z_{ij},\psi)=\prod_{i=۱}^{k}[\frac{۱}{\pi\sigma_i^{۲}\sqrt{۱-\delta^{۲}(\alpha_i)}}$$
$$\times \exp\{-\frac{۱}{۲\sigma_i^{۲}(۱-\delta^{۲}(\alpha_i))}[(x_j-\mu_i)-۲\delta(\alpha_i)$$
$$\times (x_j-\mu_i)t_j+t_j^{۲}]\}]^{z_{ij}}$$

بنابراین می توانیم تابع درستنمایی را برای داده های کامل به فرم زیر بنویسیم:

$$L(\psi|\underline{x},\underline{t},\underline{z})=\prod_{j=۱}^{n}\prod_{i=۱}^{k}[\frac{۱}{\pi\sigma_i^{۲}\sqrt{۱-\delta^{۲}(\alpha_i)}}$$
$$\times\exp\{-\frac{۱}{۲\sigma_i^{۲}(۱-\delta^{۲}(\alpha_i))}[(x_j-\mu_i)$$
$$-۲\delta(\alpha_i)(x_j-\mu_i)t_j+t_j^{۲}]\}]^{z_{ij}}\prod_{j=۱}^{n}\prod_{i=۱}^{k}\pi_i^{z_{ij}}$$

[۹] Hierarchical Model

لگاریتم تابع درستنمایی فوق را به دست می آوریم، می‌توان نوشت:

$$\begin{aligned} l(\psi|\underline{x},\underline{t},\underline{z}) &= \sum_{j=1}^{n}\sum_{i=1}^{k} Z_{ij}\{log(\pi_i) - log(\pi\sigma_i^2) \\ &- \frac{1}{2}log(1-\delta^2(\alpha_i)) \\ &- \frac{t_j^2 - 2\delta(\alpha_i)(x_j-\mu_i)+(x_j-\mu_i)^2}{2\sigma_i^2(1-\delta^2(\alpha_i))} \end{aligned} \quad (۲۳)$$

۶ برآوردهای ماکسیمم درستنمایی پارامترها

با مشتق گیری از لگاریتم تابع درستنمایی (۲۳) نسبت به π_i، μ_i، σ_i^2 و α_i هاو مساوی صفر قرار دادن معادلات به دست آمده، داریم:

$$\pi_i = \frac{\sum_{j=1}^{n} Z_{kj}}{n}، \quad \mu_i = \frac{\sum_{j=1}^{n} Z_{ij}X_j - \delta^{-1}(\alpha)\sum_{j=1}^{n} Z_{ij}T_j}{\sum_{j=1}^{n} Z_{ij}}،$$

$$\sigma_i^2 = \frac{\sum_{j=1}^{n} Z_{ij}T_j^2 - 2\delta^{-1}(\alpha_i)\sum_{j=1}^{n} Z_{ij}T_j(X_j-\mu_i)+\sum_{j=1}^{n} Z_{ij}(X_j-\mu_i)^2}{2(1-\delta^2(\alpha_i))\sum_{j=1}^{n} Z_{ij}}$$

و

$$\begin{aligned} &\sigma_i^2\delta(\alpha_i)(1-\delta^2(\alpha_i))\sum_{j=1}^{n} Z_{ij} \\ &+ (1+\delta^2(\alpha_i))\sum_{j=1}^{n} Z_{ij}T_j(X_j-\mu_i) \\ &- \delta(\alpha_i)\sum_{j=1}^{n} Z_{ij}T_j^2 - \delta(\alpha_i)\sum_{j=1}^{n} Z_{ij}(X_j-\mu_i)^2 = 0 \end{aligned}$$

حل معادله آخر می تواند مقدار$\delta(\alpha_i)$ و در نتیجه مقدار α_i را در اختیار ما قرار دهد.

همان گونه که مشاهده می‌شود در تمام معادلات فوق Z_{ij} و T_j ها ظاهر شده‌اند که عملاً مقادیر نامعلوم بوده و باید پیش بینی شوند. لذا برای پیش بینی آنها از الگوریتم EM— استفاده می‌کنیم. با در نظر گرفتن مقادیر $\mu_i^{(k)}$ ، $\sigma_i^{2(k)}$ و $\pi_i^{(k)}$ و $\alpha_i^{(k)}$ به عنوان مقادیر پارامترهای μ_i، σ_i^2، α_i و π_i در مرحله k —ام مقادیر Z_{ij} و T_j ها را در مرحله k —ام پیش بینی می‌کنیم.

$$\begin{aligned} \hat{Z}_{ij} &= E_{\hat{\psi}}(Z_{ij}|\underline{x}) \\ &= \frac{\hat{\pi}_i^{(k)} g(x_j|\hat{\mu}_i^{(k)},\hat{\sigma}_i^{2(k)},\hat{\alpha}_i^{(k)})}{\sum_{i=1}^{k}\hat{\pi}_i^{(k)} g(x_j|\hat{\mu}_i^{(k)},\hat{\sigma}_i^{2(k)},\hat{\alpha}_i^{(k)})} \end{aligned} \quad (۲۴)$$

$$\begin{aligned} \hat{S}_{2ij}^{(k)} &= E_{\hat{\psi}}(Z_{ij}T_j^2|\underline{x}) = \hat{Z}_{ij}^{(K)^2}[\hat{\mu}_{T_{ij}}^{(K)^2} + \hat{\sigma}_{T_i}^{2(K)} \\ &\times \frac{\phi\{\hat{\alpha}^{(k)}(\frac{x_j-\hat{\mu}_i^{(k)}}{\hat{\sigma}_i^{(k)}})\}}{\Phi\{\hat{\alpha}^{(k)}(\frac{x_j-\hat{\mu}_i^{(k)}}{\hat{\sigma}_i^{(k)}})\}}\hat{\mu}_{T_{ij}}^{(K)}\hat{\sigma}_{T_i}^{(K)}] \end{aligned} \quad (۲۵)$$

بنابراین با پیش بینی مقادیر Z_{ij} و T_j ها، اکنون می‌توانیم مقادیر برآورد پارامترها را در مرحله $(k+1)$ —ام به دست آورده و الگوریتم EM— را ادامه داده تا به همگرایی برآورد پارامترها برسیم.

اگر مقدارپیش بینی $\hat{Z}_{ij}$ و مقدارپیش بینی $\hat{T}_j$ باشد، در این صورت برآورد ماکسیمم درستنمایی برای π_i، μ_i، σ_i^2 و α_i به صورت زیر خواهد بود.

$$\hat{\pi}_i = \frac{\sum_{j=1}^{n}\hat{Z}_{kj}}{n}; i = 1,2,...,k \quad (۲۶)$$

$$\hat{\mu}_i = \frac{\sum_{j=1}^{n}\hat{Z}_{ij}X_j - \delta^{-1}(\hat{\alpha})\sum_{j=1}^{n}\hat{Z}_{ij}\hat{T}_j}{\sum_{j=1}^{n}\hat{Z}_{ij}} \quad (۲۷)$$

$$\begin{aligned} \hat{\sigma}_i^2 &= \frac{1}{2(1-\delta^2(\hat{\alpha}_i))\sum_{j=1}^{n}\hat{Z}_{ij}}\{\sum_{j=1}^{n}\hat{Z}_{ij}\hat{T}_j^2 \\ &- 2\delta^{-1}(\hat{\alpha}_i)\sum_{j=1}^{n}\hat{Z}_{ij}\hat{T}_j(X_j-\hat{\mu}_i) \\ &+ \sum_{j=1}^{n}\hat{Z}_{ij}(X_j-\hat{\mu}_i)^2\} \end{aligned} \quad (۲۸)$$

با استفاده از الگوریتم $EM-$ و جایگزینی $\hat{S}_{1ij}^{(k)}$ و $\hat{S}_{2ij}^{(k)}$ به جای $\hat{T}_j$ و $\hat{T}_j^2$ می‌توانیم برآورد پارامترها را در مرحله $(k+1)$ — ام به صورت زیر به دست آوریم.

$$\hat{\pi}_i^{(k+1)} = \frac{\sum_{j=1}^{n} \hat{Z}_{kj}^{(k)}}{n}; i = 1, 2, ..., k \qquad (29)$$

$$\hat{\mu}_i^{(k+1)} = \frac{\sum_{j=1}^{n} \hat{Z}_{ij}^{(k)} X_j - \delta^{-1}(\hat{\alpha}_i^{(k)}) \sum_{j=1}^{n} \hat{S}_{1ij}^{(k)}}{\sum_{j=1}^{n} \hat{Z}_{ij}^{(k)}} \qquad (30)$$

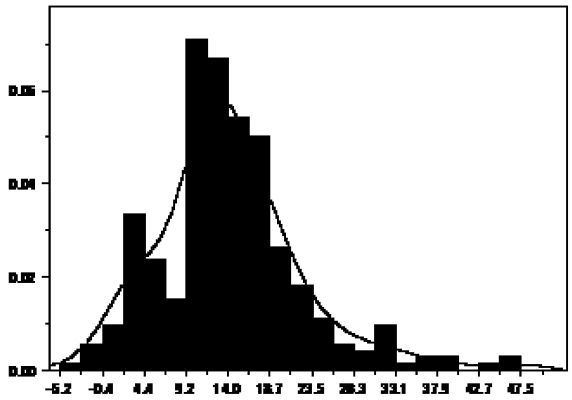
$$\begin{aligned} \hat{\sigma}_i^{2(k+1)} &= \frac{1}{2(1-\delta^2(\hat{\alpha}_i^{(k)})) \sum_{j=1}^{n} \hat{Z}_{ij}^{(k)}} \{\sum_{j=1}^{n} \hat{S}_{2ij}^{(k)} \\ &- 2\delta^{-1}(\hat{\alpha}_i^{(k)}) \sum_{j=1}^{n} \hat{S}_{1ij}^{(k)} (X_j - \hat{\mu}_i^{(k)}) \\ &+ \sum_{j=1}^{n} \hat{Z}_{ij}^{(k)} (X_j - \hat{\mu}_i^{(k)})^2 \} \end{aligned} \qquad (31)$$

با ثابت نگه داشتن $\mu_i = \hat{\mu}_i^{(k+1)}$ و $\sigma_i^2 = \hat{\sigma}_i^{(k+1)}$، در معادله زیر، $\hat{\alpha}_i^{(k+1)}$ را به عنوان برآورد پارامتر α_i $(i = 1, 2..., k)$ به دست می‌آوریم.

$$\begin{aligned} &\sigma_i^2 \delta(\alpha_i)(1-\delta^2(\alpha_i)) \sum_{j=1}^{n} Z_{ij} + (1+\delta^2(\alpha_i)) \\ &\times \sum_{j=1}^{n} Z_{ij} T_j (X_j - \mu_i) - \delta(\alpha_i) \sum_{j=1}^{n} Z_{ij} T_j^2 \\ &- \delta(\alpha_i) \sum_{j=1}^{n} Z_{ij} (X_j - \mu_i)^2 = 0 \end{aligned}$$

در اینجا مثالی را در مورد آمیخته دو چگالی نرمال چوله بررسی می‌کنیم.

مثال ۶—۱ با به کارگیری قضیه (۱) تعداد ۳۰۰ داده را از دو توزیع نرمال چوله آمیخته با پارامترهای $\mu_1 = 5$، $\mu_2 = 10$ و $\sigma_1 = 5$، $\sigma_2 = 10$ و پارامترهای شکل $\alpha_1 = 0/5$ و $\alpha_2 = 2$ با نسبتهای ۰/۳ و ۰/۷ تولید کرده‌ایم. می‌خواهیم برآورد پارامترها را به دست آوریم. ابتدا هیستوگرام داده‌ها را رسم می‌کنیم. چولگی و آمیختگی داده‌ها کاملا مشخص است.



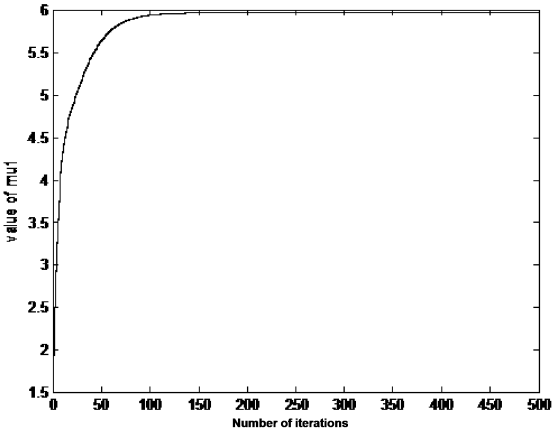
شکل ۴. هیستوگرام داده‌های مربوط به چگالی آمیخته از دو توزیع نرمال چوله.

با نوشتن برنامه مربوط به الگوریتم $EM-$ در نرم افزار $MATLAB$ و اجرای این برنامه، برآورد پارامترها در جدول زیر به دست آمده است.

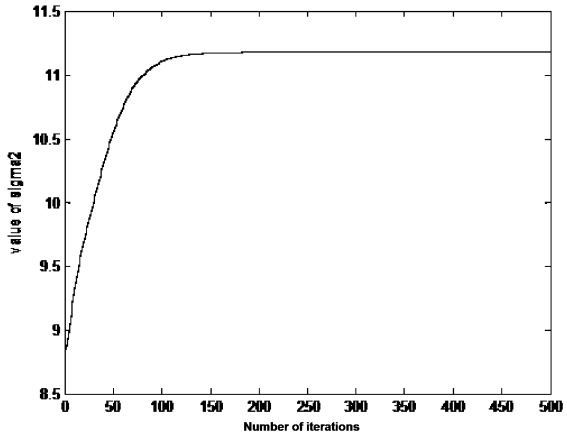
جدول ۱. مقادیر برآورد پارامترهای چگالی آمیخته از دو چگالی نرمال چوله.

$Parameter$	$True\ value$	$Estimation$	$Std.$
μ_1	۵	۵/۳۶۶۲	۰/۱۵۸۶
μ_2	۱۰	۹/۹۶۴۱	۰/۰۳۰۵
σ_1	۵	۴/۸۲۳۰	۰/۰۸۴۹
σ_2	۱۰	۱۱/۱۷۶۹	۰/۹۹۳۱
α_1	۰/۵	۰/۵۲۸۴	۰/۰۷۹۷
α_2	۲	۱/۹۶۸۴	۰/۰۲۷۴
π	۰/۳	۰/۲۷۴۷	۰/۰۰۶۶

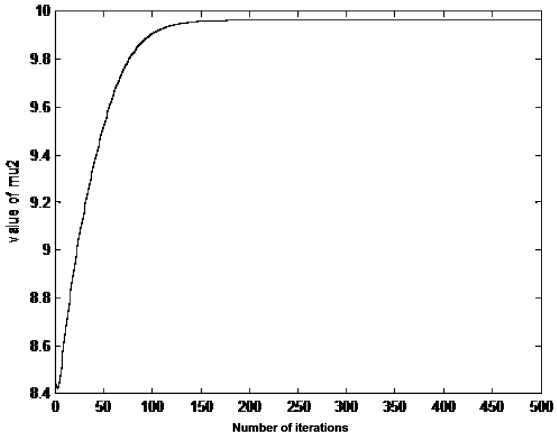
برای اطمینان از درست بودن برآوردهای به دست آمده نمودارهای همگرایی برآورد پارامترها را رسم می کنیم.



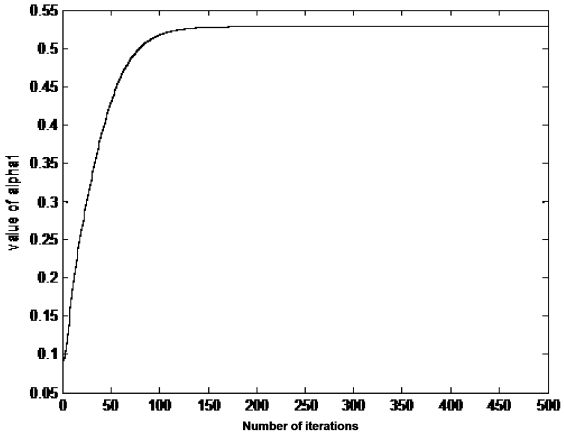
شکل ۵. نمودار همگرایی برآورد پارامتر μ_1 .



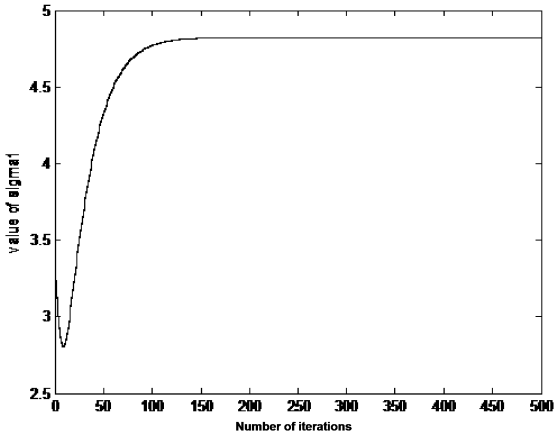
شکل ۸. نمودار همگرایی برآورد پارامتر σ_2 .



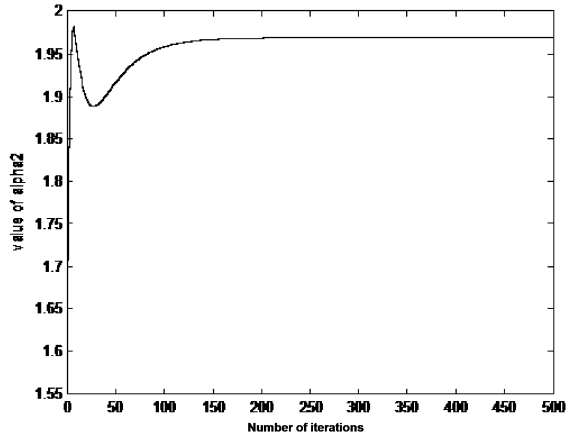
شکل ٦. نمودار همگرایی برآورد پارامتر μ_2 .



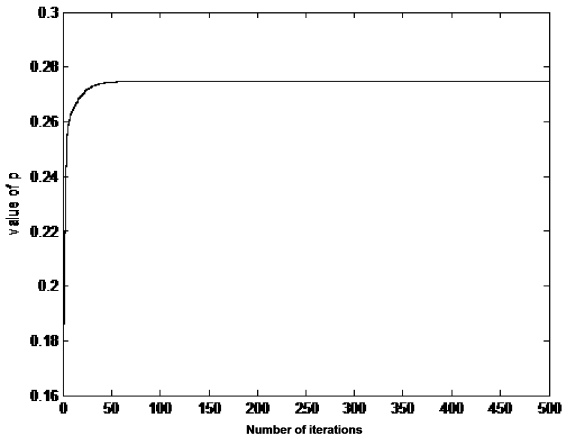
شکل ۹. نمودار همگرایی برآورد پارامتر α_1 .



شکل ۷. نمودار همگرایی برآورد پارامتر σ_1 .



شکل ۱۰. نمودار همگرایی برآورد پارامتر α_2 .



شکل ۱۱. نمودار همگرایی برآورد پارامتر π .

مراجع

[1] Azzalini, A. (1986). Further results on a class of distributions which includes the normal ones, *Statistica*, 46,199-208.

[2] Dempster, A. P., Laird, N. M and Rubin D. (1977). Maximum Likelihood from incomplete Data via the EM algorithm, *Jornal of Royal Statistical Society*, B, 39, 1-38.

[3] Reinaldo, B. (2003). *A New Class of Skew-Normal Distributions*, Technical Report, North Carolina State University, Institue of Statistics, 1-15.

[4] Pewsey, A. (2001). Problems of inference for Azzalini's skew-normal distribution, *Journal of Applied Statistics*, 27, 859-870.